

BESZÉDKUTATÁS 2014

**Szerkesztette:
Gósy Mária**

MTA Nyelvtudományi Intézet
Kempelen Farkas Beszédkutató Laboratórium
Budapest, 2014

A szerkesztőbizottság elnöke: Gósy Mária (MTA Nyelvtudományi Intézet)
A szerkesztőbizottság tagjai: Bunta Ferenc (University of Houston)
Horváth Viktória (MTA Nyelvtudományi Intézet)
Mády Katalin (MTA Nyelvtudományi Intézet)
Markó Alexandra (Eötvös Loránd Tudományegyetem)
Olaszy Gábor (Budapesti Műszaki és Gazdaságtudományi Egyetem)
Siptár Péter (Eötvös Loránd Tudományegyetem)
Vago, Robert (Queens College és City University of New York)

A korpusz szerkesztésében részt vettek: Németh Mátyás
Bóna Judit
Váradi Viola
Imre Angéla
Markó Alexandra

Az egyes tanulmányokat szakterületi kompetenciával rendelkező szakemberek lektorálták.

Technikai szerkesztés: Gráci Tekla Etelka
Technikai munkatársak: Markó Alexandra, Neuberger Tilda, Váradi Viola

A kötet kiadását az MTA Könyv- és Folyóíratkiadó Bizottsága támogatta.

ISSN 1218-8727

© MTA Nyelvtudományi Intézet

Felelős kiadó: Kenesei István igazgató

Készült: Regiszter Kiadó és Nyomda Kft.

TARTALOM

Kiss Jenő: Spontán beszéd, szövegalkotás, nyelvjárások, dialektológia	5
Gósy Mária: A palatális közelítőhang kétféle funkcióban	17
Gráci Tekla Etelka: Magánhangzó-időtartam a követő mássalhangzó zöngésségének függvényében	41
Neuberger Tilda: Zöngétlen explozívák időviszonyai a beszéd típus és az egyéni sajátosságok tükrében	56
Markó Alexandra: Az irreguláris zöngeminőség gyakorisága és pozíciói különféle spontán beszédhelyzetekben	69
Horváth Viktória: Szóidőtartamok gyermekek és felnőttek spontán narratíváiban	87
Tóth Andrea: Gyermekek nemének és életkorának meghatározása a beszédük alapján	98
Szántó Anna: A célzott beszédpercepció- és beszédértés-fejlesztés hatása utánkövetéses vizsgálatban	112
Bóna Judit: Megakadásjelenségek az életkor, a nem és a beszéd típus függvényében	123
Gocsál Ákos: Az artikulációs tempó vizsgálata régi magyar filmekben	144
Deme Andrea: A mássalhangzó-környezet és a zöngéindulás hatása az énekelt magánhangzók észlelésére	155
Beke András: Egyszerre beszélések detektálása a beszélődetektálás javításához	172
Csapó Tamás Gábor – Németh Géza: Irreguláris beszéd regulárisá alakítása beszéd kódoláson alapuló módszerrel	193
Olasz Gábor – Abari Kálmán – Bartalis Máttyás: Magyar hangsúlyjelölési szöveges adatbázis fejlesztése és referencia-vizsgálata	205
Kohári Anna: Tempóváltozások a vizsgált szakasz nagyságának függvényében	220
Mátyus Kinga – Orosz György: MONYÉK – Morfológiailag egyértelműsített óvodai nyelvi korpusz	237
Bóna Judit – Imre Angéla – Markó Alexandra – Váradi Viola – Gósy Mária: GABI – Gyermeknyelvi Beszédatadabázis és Információtár	246

„Nyelvbotlás”-korpusz.....	253
1. Téves szó.....	254
2. Grammatikai hiba.....	258
3. Kontamináció.....	261
4. Téves kezdés.....	262
5. „Nyelvem hegyén van” jelenség.....	264
6. Perszeveráció.....	265
7. Anticipáció.....	267
8. Metatézis.....	268
9. Egyszerű nyelvbotlás.....	270
10. Több típusba sorolható jelenségek.....	271
A magyar és a nemzetközi tudományos élet eseményei.....	277
Björn E. Lindblom 80 éves • <i>Deme Andrea</i>	277
A kötet szerzői.....	282

SPONTÁN BESZÉD, SZÖVEGALKOTÁS, NYELVJÁRÁSOK, DIALEKTOLÓGIA

Kiss Jenő

Bevezetés

A következő témakörök kerülnek szóba: 1. A nyelvjárási szövegek és vizsgálatuk. 2. Melyek a főbb nyelvjárási szövegalkotási jegyek? 3. Van-e a nyelvjárási szövegalkotás vizsgálatának a dialektológián túlmutató szerepe?

Szándékosan markolok keveset. Csupán alkalmazom azon szövegtani standardismereteket, amelyek tanulmányom célkitűzésének megvalósításához itt és most elegendőnek látszanak. Ennek megfelelően szabtam szűkre a szakirodalmi hivatkozásokat is.

Mikor spontán a beszéd? Akkor, „ha a beszélő mindenféle előzetes tervezés és felkészülés nélkül az adott helyzetben fogalmazza meg gondolatait, rendeli hozzájuk a megfelelő nyelvi formát, és hangosítja a közléseket. A beszéd spontaneitása nem függ attól, hogy a beszéd narratíva-e, vagy társalgásban hangzik el. A spontaneitás független a beszélő aktuális lelkiállapotától vagy az ismereteitől” (Gósy et al. 2012: 13). A spontán beszéd legfőbb jellemzői pedig a gyakori közbeékelések, a befejezetlen mondatok, az alanyváltások, az egyeztetési következtelenségek (vö. még Szende 1973: 25; Szabó 1983; Keszler 1983: 190; Szikszainé Nagy 1999: 377; Lanstyák 2009: 16).

A nyelvjárási szövegek és vizsgálatuk

A nyelvjárási szövegkutatásra időben az első lehetőséget a nyelvjárási szöveggyűjtemények adták. Az első 110 évvel ezelőtt jelent meg Simonyi Zsigmond jóvoltából, aki 1902-ben tette közzé tréfás népmeséket és adomákat tartalmazó nyelvjárási olvasókönyvét (a sor legújabb tagja Silling István 2010-es kötete: *Vajdasági magyar nyelvjárási olvasókönyv*). A néprajzosok, folkloristák, mesekutatók nevéhez számos olyan szöveggyűjtemény kiadása fűződik, amelyek többé-kevésbé úgyszintén alkalmasak bizonyos regionális szövegalkotási sajátságok vizsgálatára. Nyelvészeti szempontból a leghitelesebb források azonban a spontán beszéd produktumai, illetőleg a hangrögzítések korpuszok. Ez utóbbiak azonban technikai fejlettséghez kötődnek, s nálunk nem régebbiek az 1940-es éveknél. Az természetes, hogy a modern textológia megszületése előtt a nyelvjárási szövegek sem lehettek korszerű szövegnyelvészeti vizsgálat tárgyai. Most azonban más a helyzet. A technikai lehetőségek és a vizsgálati korpusz lényegében adva van. A vállalkozó szellemű kutatók hiánya a legnagyobb gond.

Nyilvánvaló, hogy a nyelvjárási spontán beszéd vizsgálatára a múltra vonatkozóan a nyelvjárási hangfelvételek kínálják a legjobb, a kutatótól azonban minden esetben kellő körültekintést kívánó lehetőséget. A magyar nyelvjárási hangfelvételek sora (nem tekintve a szórványos, technikailag még kezdetleges fonográffelvételeket) 1940-ben indult Hegedűs Lajosnak köszönhetően (ezek a felvételek a Nyelvtudományi Intézet hangarchívumában vannak, feldolgozásukat fonetikus szakértők végzik). A sorban a nagyatlasz (*A magyar nyelvjárások atlasza*) ellenőrző gyűjtései során 1960 és 1964 között készült szövegfelvételek következnek, mert ezek között vannak spontánbeszéd-felvételek is (l. Végh–Balogh 1975). Illetőleg az *Új magyar nyelvjárási atlasz* szövegfelvételei, amelyeket a gyűjtők 2008 és 2011 között rögzítettek diktafonnal. Mindkét atlasz szövegfelvételei digitalizálva vannak: az előbbieket a Nyelvtudományi Intézetben és az ELTE Magyar Nyelvtudományi és Finnugor Intézetében, az utóbbiak az említett egyetemi intézetben vannak. 1982 májusában Balogh Lajos mint témafelelős az Imre Samuval történt tanácskozás alapján a Nyelvtudományi Intézetben készített egy tervezetet a nyelvjárási és a regionális köznyelvi szövegek magnetofonnal történő gyűjtéséről. Az Intézet a gyűjtésben való részvételre és a szervezési és koordináló feladatok ellátására volt kijelölve, az egyetemek és főiskolák (szám szerint tíz) magyar (nyelvészeti) tanszékeire általában a többé-kevésbé vonzáskörükbe tartozó településeken való gyűjtés várt. A szándék az volt, hogy a nagyatlasz 20 évvel korábban készült magnetofonos felvételei után korszerűbb technikai eszközökkel készüljenek hangfelvételek azonos szempontok szerint. Az elkészült gépirásos dokumentum a *Magnetofonnal történő szöveggyűjtés a területi nyelvváltozatok köréből* címet viseli. A felvételek készítését – tekintettel a magasabb igényekre – elsősorban szakképzett kutatóktól várták, de be lehetett vonni gyakorlottabb egyetemi, főiskolai hallgatókat, a tudományos diákkörök tagjait is a munkába (Papp Lajos 1983: 1; l. még Balogh Lajos 2014). 1986-tól négy éven át az Országos Tudományos Kutatási Alap támogatásával folytak a kutatások, s az első szakaszban anyaggyűjtés folyt, ti. szövegfelvételek készültek, elsősorban vidéki egyetemek és főiskolák tanárainak köszönhetően: ebből a gyűjtésből készült is egy szöveggyűjtemény, amely 1993-ban jelent meg (Balogh 1993: 7). A szóban forgó szövegfelvételek egy-egy digitalizált példánya megvan a Nyelvtudományi Intézetben és az ELTE említett intézetében (Balogh Lajos, Posgay Ildikó és Bodó Csanád tájékoztatása szerint). Az intézeti munkálattal összefüggésben indult az ELTE-n G. Varga Györgyi irányításával annak a diák-kutatócsoportnak a tevékenysége, amelynek során regionális köznyelvi beszédfelvételek is készültek. Ezeket P. Lakatos Iлона az intézeti hangarchívumnak adta át, tehát ezek is megvannak, s már elkezdődött archiválásuk, azaz modern adathordozókra történő átvitelük (Gósy Mária szíves közlése). Megemlítendő azoknak a kutatócsoportoknak a gyűjtési tevékenysége is, amelyeknek köszönhetően nagy mennyiségű regionális nyelvi anyagot, így szövegfelvételeket is rögzítettek (Magyarországon a

budapesti bölcsészkar mellett a piliscsabai, a veszprémi, illetőleg a nyíregyházi és szombathelyi dialektológusok nyelvjárasi és regionális köznyelvi gyűjtéseire gondolhatunk elsősorban). De nem feledkezhetünk meg azokról a nyelvjárasi szövegfelvételekről sem, amelyek nem valamely központi kutatás vagy pályázati projekt keretében születtek. Az egyetemi és főiskolai magyar szakos hallgatók tanári vezetéssel szervezett és irányított nyelvjárásgyűjtéseken készített regionális szövegfelvételeiről, s végül, de nem utolsósorban a dialektológusok saját regionális nyelvi fonotékáiról van szó. Bizonyos, hogy az idetartozó felvételek együttesen igen nagy beszélt nyelvi korpuszt jelentenek. Ezeknek a mennyiségéről, minőségéről, felhasználhatóságáról, területi eloszlásáról, felvételük idejéről, jelenlegi állapotukról sajnos nincs pontos tudomásunk. Hitelesen csak az ELTE-vel kapcsolatban nyilatkozhatom, ahol volt egy ígéretes terv: „Az ELTE-n készülő Dialektológiai Archívum létrehozásával [...] az a cél, hogy az összegyűlt nyelvjárasi hanganyagok korszerű hordozókra kerüljenek át, lehetővé téve ezzel a mai igényeknek megfelelő használatot, egyszersmind az adatok kellő minőségben való átmentését” (Kiss 2003: 142). Szerettünk volna olyan archívumot létrehozni, amely be tudta volna fogadni más intézmények gyűjtéseit is. (Részletesebben l. Radványi 1997–1999; Radványi–Vékás 1999; Vékás 1999a, 1999b; vö. még Cs. Nagy 2013: 72.) Ebben Radványi Péternek volt különösen technikai oldalról fontos szerepe. Vékás Domokos úttörő dialektológiai szoftverfejlesztő tevékenységével új távlatok és új lehetőségek nyíltak, a korábbi tervnek korszerűsített formában való megvalósítására pedig azért nem került sor, mert egyrészt az új nyelvtechnológia megjelenésével az anyagfeldolgozás igénye került előtérbe, másrészt és bizonyára leginkább azért, mert a kivitelezéshez (a nagy mennyiségű szövegfelvétel modern adathordozókra viteléhez) szükséges anyagiakat – pályázati és egyéb támogatások hiányában – nem sikerült előteremteni, illetőleg mert nem volt kellő motiváltság azokban, akik a megvalósításban tevélegesen közreműködhetek volna. Aligha vitatható: a feladat elvégzése sürgető, sőt egyre sürgetőbb (a felvételek technikai minőségének elavulása miatt; l. Kontra 2013). Am ami központilag nem megy, tudniillik az adatátmentés, azt az egyes intézmények saját hatáskörükben el tudnák végeztetni, ha a személyi és tárgyi feltételeket elő tudják teremteni. Előttünk a nagyatlász szövegfelvételeinek a sorsa áll a végül is megnyugtató végkifejlettel. Ugyanis a nagyatlász nyelvjárasi szövegfelvételek, amelyek a Nyelvtudományi Intézetben voltak, az állapotukra vonatkozó vészjósló híradások ellenére – legalábbis átmenetileg – „révbe értek”. Ugyanis 2004-ben sikerült pályázati támogatással több részfeladatot is elvégeznünk, köztük azt is, amellyel egyetlen célunk a nyelvatlaszos hangfelvételek megmentése volt. A Nyelvtudományi Intézet igazgatója, Kenesei István megkeresésemre készségesen hozzájárulását adta az Intézet tulajdonában lévő korpuszsal való munkánkhoz, s fölkeresemre, konzorciumi keretben a Nyelvtudományi Intézet munkatársa, Posgay Ildikó gondoskodott a szalagokon tárolt anyag digitalizálásának elvé-

geztetéséről. Ami a digitalizálást illeti, e tekintetben a nagyatlász és a Hajdú–Kázmér-féle nyelvjárási szöveggyűjtemény szövegei, illetve egy részük van digitalizálva azon ELTE-s projekteknek s elsősorban Vargha Fruzsínának köszönhetően, amelyek a Vékás Domokos nevével fémjelzett dialektológiai szoftverfejlesztéssel vannak összefüggésben (l. még Bárh–Vargha 2011).

A Nyelvtudományi Intézet hangarchívuma is őrzi nyelvjárási szövegfelvételeket, így a Hegedűs Lajos-féle régi, nyelvjárástörténetileg is értékes felvételek az Intézet gondozásában vannak. Ezek a felvételek megfelelő módon vannak digitalizálva és tárolva (az intézeti internetes nyelvjárási hangarchívumra l. Nikléczy–Horváth 2007). Idekapcsolódó apró tudománytörténeti megjegyzés: Szépe György megőrzésre hosszú évekkel ezelőtt Pesti Jánosnak adott át Hegedűs-féle nyelvjárási hangfelvételeket, amelyek bakelitlemeze voltak rögzítve. Ezeket Pesti János néhány éve elküldte nekem, én pedig az intézeti hangarchívumnak adtam tovább. Mondhatni, szerencsés körülmények folytán kerültek ezek a felvételek a megfelelő helyre. Ez azt juttatja eszembe, bizony kíváncsi volna, hogy a tanszékeken itt-ott található regionális szövegfelvételek megőrzése ne a pusztá véletlenül, ne is csupán egy-egy őket később is értékelni tudó nyelvész vagy nem nyelvész kolléga jóindulatán múljon, hanem időben végrehajtott, céltudatos „mentőakció”-k biztosítsák továbbélésüket.

Bár a korábbi, a magnetofon előtti korszak nyelvjárási szövegjegyzéseivel szemben nem ok nélkül vagyunk szkeptikusak, kár volna a fürdővízzel együtt a gyereket is kiöntetni. Arra gondolok, hogy bizonyos nyelvjárási szövegyszerkesztési sajátosságok – elsősorban akkor, ha a gyűjtő anyanyelvjárási szövegeit tette közzé – általában kellően hitelesnek tekinthetők. Leginkább arról lehet szó, hogy archaikusabb állapotot tükröznek, mivel korábban általános szempont volt, hogy a nyelvjárási különlegességeket, ritkaságokat, a köznyelvtől való minél nagyobb eltéréseket rögzítsék, s ez vonatkozott a szövegformálásra és a stílusra is. Tehát a Bárczitól említett „jóhiszemű hamisítványok” kitétel nem a szövegalkotásra vonatkozhatott elsősorban, hanem a hang- és alaktani „megoldások”-ra.

Főbb nyelvjárási szövegalkotási jegyek a korábbi kutatások tükrében

Nyelvjárási szövegalkotással behatóan nem foglalkozott még senki sem a magyar nyelvtudományban. De a témakör több tanulmánynak is tárgya így vagy úgy (Kiss 1982; Vincze 1986a, 1986b; Büky 1989–1990; Szabó 1994; vö. még Hegedűs 2001: 362–363). Persze amíg nem volt a beszélt nyelvet is vizsgálata tárgyának tekintő szövegtani kutatás, addig nyelvjárási szövegtani vizsgálat sem lehetett. Mindenütt így volt ez. Az alapos, precíz kézikönyv, a *Dialektologie. Handbuch der deutschen und allgemeinen Dialektforschung* (DialHandb. 1982–1983) két vaskos kötetében sem találunk olyan alfejezetet sem, amelynek tárgya a nyelvjárási szövegalkotás volna. Annak, hogy a nyelvjáráskutatás nem vett fordulatot újabban sem a szövegtan irányába

(ahogy azt Szabó Zoltán 1994-es tanulmányában megfogalmazta: „tulajdonképpen a >>nyelvjárási textológia<< kellene hogy a nyelvjáráskutatás legjelentősebb ága legyen”, 58), kettős oka van. Egyrészt az, hogy a dialektológiát saját belső logikája a korszerűsödés más irányába, tudniillik a változásvizsgálatnak és a szociodialektológia kialakításának az irányába vitte. Merthogy nem a nyelvi szerkezet, hanem a nyelvjárási nyelvhasználat társadalmi meghatározottsága került a dialektológiai vizsgálatok fókuszába. Másrészt pedig az a tény, hogy a nyelvjárástani kutatások középpontjában hagyományosan az alsóbb nyelvi szintek vizsgálata állt mindig is – nagyrészt azért, mert a magasabb nyelvi szinteken, így már a mondatszinten is kevés és csekély különbség van a magyarban a köznyelv és a nyelvjárások között (l. Szabó 1983: 52; Kiss 1982: 15–16). Köztudott ugyanis, hogy minden nyelvváltozat – a szókészlet kivételével – „általában ugyanazt az általános mintát követi” (Tolcsvai Nagy 1996a: 56).

Kézenfekvő, hogy a nyelvjárási szövegalkotás a beszélt nyelv, sőt a spontán beszélt nyelv jellemzőinek a megismerésére jó lehetőséget kínál. Hiszen – a magyar nyelvterületen legalábbis – a nyelvjárások lényegében csak beszélt formában, és pedig a hétköznapi spontán beszélt nyelviségnek, a tervezetlen nyelvi diskurzusoknak a formájában élnek. A spontán beszélt nyelvjárási szövegektől a legtávolabb természeténél fogva az írott standard áll a maga szépirodalmi és tudományos szövegeivel. Ezért az írott nyelv és a spontán nyelvjárási nyelvhasználat közötti különbségek a legnagyobbak, s elsődlegesen a társadalmi (szociokulturális) környezet eltérő voltában lelik eredendő magyarázatukat. Abban tudniillik, hogy a nyelvjárási beszélők nagy többségének a rendszeres írástevékenység, különösen az írott nyelvi szövegalkotás nem mindennapi kenyeré. Természetes tehát, hogy a nagyobb részt a szóbeliségen alapuló gondolkodás és kifejezőmód jóval jellemzőbb a nyelvjárási beszélőkre és spontán nyelvhasználatukra, mint az írásbeliség világában is nap mint nap aktívan részt vevő köznyelvi, iskolázott beszélőkre és nyelvhasználatukra.

Még világosabbá válik e gondolatsor akkor, ha hivatkozom arra, amit Walter J. Ong egyik, szempontunkból fontos munkájában, a *Szóbeliség és írásbeliség* címűben a nagyobb részt a szóbeliség világában élő beszélők nyelvhasználatának jellemző jegyeiként sorol föl: inkább mellérendelő, mint alárendelő; inkább halmozó, mint elemző; redundáns; konzervatív, tradicionális; az emberi életvilághoz közelálló; inkább empátikus és közvetlen, mint objektív és távolságtartó; homeosztatis, azaz egyensúlyban tartó (a jelen számára semmilyen relevanciával nem bíró emlékektől igyekeznek a beszélők megszabadulni); inkább szituációfüggő, mint elvont (2010: 38–55). Ochs (2010) a tervezetlen beszélt nyelvi diskurzusra – ilyenek többnyire a nyelvjárási megnyilatkozások – jellemzőnek mondja a következőket: a deiktikus elemek túlsúlyban vannak; kerülnek az alárendelő mondatokat; gyakori a párhuzamoság, az ismétlések (lexikai elemek, hasonló szintaktikai szerkezetek ismétlé-

se); a narratív formákban gyakori az indító múlt idő után a jelenre való váltás (l. Andó 2006: 144). Lanstyák István a magyar spontán beszélt nyelv jellemzőiként említi az előkészítés nélkülséget, a kötetlenséget, más szavakkal a szövegfüggetlen szövegalkotást, a beszédhelyzetbe ágyazottságot, azaz azt a tényt, hogy a beszélő „ebből következően mondanivalójának számos részletét nem fejezi ki verbálisan” (2009: 16), illetőleg a szupraszegmentális eszközök és a testbeszéd kínálta lehetőségek kihasználását (17). A részletekre vonatkozóan számos releváns következtetés olvasható Keszler Borbála 1983-as tanulmányában.

A nyelvjárási szövegalkotás főbb jellemzői

A továbbiakban a leggyakrabban előforduló azon szövegalkotási sajátosságokat említem egy-két példával szemléltetve, amelyek magyar nyelvjárási oldalról is alátámasztják azt, amit idevonatkozóan Ong (2010), Ochs (1979), Gósy (2004: 228–241), Lanstyák (2009), illetőleg 1983-as tanulmányában Keszler írt.

1. Gyakoriak a közbeszúrt mondatok, kitérések, különösen az elbeszélő szövegekben (ha külön nem jelzem a forrást, akkor a példák szülőfalumból, Mihályiból való és saját gyűjtésűek):

(1) Mikor leszereltünk, azután odamentünk, mert ott volt legkésőbb zárás – **tudod, Husztról mikor jöttünk haza** – akkor találkoztam azzal is.

(2) Miki bátyám azt mondta – **ami igaz is** –, hogy abból a fajtából nem érdemes leásni egyet sem.

2. A köszönetnélküliség is jellemző jegye a nyelvjárási szövegalkotásnak. Móricztól idézek klasszikusan jó példát (*A boldog ember*):

(3) Otthon legénykedtem, szép életem vót, mán jányok után járogattam, sok szeretőm vót, ki elhagyott, kit én, kitül elverte, ki megcsalt; birtokot elvették, semmitlen maradtam, visszapereltem, megbuktam, akkor gondoltam, megpróbálom én mégis Pestet (idézi Sebestyén 1972: 31).

Két esetben Móricztól veszem a példát, mert hitelesen adják vissza sorai a nyelvjárási szövegszerkesztés több sajátosságát is.

3. Gyakoribb az egyenes beszéd, idézet is:

(4) Mondom a bírónak, mert tárgyalás lett: bíró úr – mondok – idefigyelen, hát én megmagyarázom magának, hogyan is volt.

(5) Ott volt egy ismeretlen tiszt is, aztán egyszer csak rám kiabál: mit ácsorog maga itt? – így e –, hát azt hittem, nem jól hallok!

4. Gyakoribb a hiányos mondszerkesztés is:

(6) Ahol nem volt nyitva az ajtó, ott föl a kapura, beugrani, aztán beengedni a népet.

(7) Aztán hát hogy kulák, kitakarodni még a faluból is.

5. A nyelvjárási szövegek empátikus, közvetlen, nem távolságtartó, emocionális jellege nem maradhat említetlen. Ez különösen a történetmondásokra jellemző. Az elbeszélő ugyanis mintegy újra beleéli magát abba a helyzetbe, amelyben volt az elbeszélte esemény történésének az idején (vö. Szabó József 1983: 40). Móricz írta a szatmári népnyelvről: „Csupa kép, hasonlat, személyítés [...] rajzanak beszédjében” (idézi Sebestyén 1972: 17). Czine Mihály pedig Móriczról a következőt: „a Móricz által teremtett nyelv [...] mindent velejéig megmutató természetű és érzékletes, emocionális jellegű, mint a tiszaháti népnyelv” (1960: 194). Jellemző erre a szövegalkotásra az indulatszós vagy érzelmi töltésű határozószós mondatindítás: „ezek lezárják az előzményt, bevezetik az új eseményt, és főleg: értelmi, érzelmi kapcsolatot teremtenek [...] a beszélő és a hallgató között” (Sebestyén 1972: 34). Sebestyén a következőket közli *A boldog emberből*: *asztán, avval, evvel, hát, hej, na, nahát, ó, osztán*). Az elbeszélő szövegekre jellemző jegyekként Mihályiból a következő kapcsoló funkciójú szavaknak, kifejezéseknek, közbeszűrásoknak, köztöszóknak a gyakori használatát jegyeztem föl: *aztom* 'úgy hiszem, úgy tudom' (= azt tudom), *asztán/azután, hát aztán, na most, na most akkor, na most aztán, illetőleg elég az hozzá, teszem azt, teszem föl* 'például', *tudja Isten, fene (se) tudja*. Vincze Krisztina csermelyi (tehát palóc nyelvjárási) vizsgálata alapján a következő gyakorisági sorrendet állította föl: *hát, ugye, azután/oszt, akkor, na/no, szóval, úgyhogy, úgy, ijen, érted-e* (1986a: 224).

6. Az érzelmileg fontosabb szavaknak, illetőleg a lélektani állítmánynak a mondat elejére helyezése, átképzeléses előadásmód, a praesens historicum keverése a múlt idővel. Móricztól veszek példát megint:

(8) Másnap karácsony. Mentem templomba. Nagyon körülvettek a legények, de az én cimboráim már házasok vótak. Orbán Gyurinak már gyereke is vót. Files Gyurkának is hasas vót a felesége. Majd a lesz az enyém is, csak még nem tudom, mejik. Körülnéztem a templomba, mék jány várja az áldást tőlem. Ott vót Váradi Piroska is, sápadt vót egy kicsit. Ezt kéne mégis megpirosítani (idézi Sebestyén 1972: 35).

7. Az elbeszélés némely sajátosságát már említettem. A történetmesélésre azért érdemes további figyelmet is szentelnünk, mert egyetértéssel említhető W. Labov narratívmodellje (1997), hogy tudniillik a személyes elbeszélés egyéni élményből kiinduló, tervezetlenül elhangzó beszédprodukció (vö. Andó 2006: 128). Az az elsősorban Labov és Waletzky nevéhez fűződő szociolingvisztikai hagyomány (1967), amely szerint a történeteket a hétköznapi kommunikáció részeként, a tárgyalás aktusába ágyazottan kell vizsgálni (vö.

Andó i. h.), a nyelvjárási szövegalkotás tekintetében bizonyosan maradéktalanul igaz. A történetmesélést előadásként is felfoghatjuk, amennyiben úgy tekintjük, hogy a beszélő egy történetet elmesélve lehetőséget ad a hallgatónak, hogy átéljék, megtapasztalják az eseményeket. Az ilyen előadásmódnak, s a praesens historicumnak kétségtelenül dinamizáló, élénkítő hatása, aktualizáló szerepe van (Wolfson után Andó 2006: 131). Egy példa az 1970-es évek végéről (köznyelvire áttéve a szavakat): „Gabinak mondom tegnap este – mondok – itt van húsz forint e – mondok – kéménypénzt is szed a kéménysprő – mondok – ha tehozzád még ma odamegy, mondd meg neki, az én nyugtámat is adja oda. Reggel elmegy a (tej)csarnokba, innen meg hazamegy, ott van a villanyos [= díjbeszedő], fizet neki, visszajön ide, jaj – azt mondja – nem adtam oda neki a villanypénzedet – azt mondja – itt van nálam e, most mit csináljak vele – azt mondja” (Kiss 1982: 75).

Megemlítendő néhány további sajátosság is: a hiányos mondatok nagyobb aránya, a gondolatmenet folyamatosságában jelentkező hiányok, a módosító határozószócskák gyakorisága, az alkalmoszerű emberi elbizonytalanodásnak a jelei, a megszakított mondatok, a módosító mondatrészletek, a főmondati hiány (ezekre l. Keszler 1983: 176, 181, 182, 202). A kommunikáció ezeknek ellenére is zökkenőmentes, mert a szituáció, az előzményeknek a beszélő és a hallgató által való ismerete, a mimika és gesztusnyelv fölöslegessé teszik bizonyos dolgoknak a verbális kifejezését (Keszler 1983: 183; Lanstyák 2009: 17). Herczeg Gyula a paraszti beszéd illogikus szerkesztésmódját emlegette Móricz *A boldog ember* című munkájának stílusát vizsgálva (Sebestyén 1972: 31). Nyilvánvaló, hogy elvárásait elsősorban a standardnyelvi, illetőleg az irodalmi nyelvi szövegalkotás szabályaihoz igazította. Illogikusnak nevezni a paraszti, illetőleg a nyelvjárási szerkesztésmódot azonban már csak azért sem lehet, mert ebben az élőbeszédes közegben ez a szerkesztésmód semmiféle kommunikációs gondot nem okoz (l. még Sebestyén 1972: 32). Az azonban, hogy meddig érdemes, célravezető elmenni népnyelvi elemek, jelenségek felhasználásában a szépirodalmi alkotásokban (Bárczi 1961), más kérdés.

A szóbeliségre jellemzőként föntebb megnevezett bizonyos szövegalkotási sajátosságok nyelvjárási közegben (saját anyanyelvjárási tapasztalataim és nyelvi anyagaim alapján is) egyértelműen gyakrabban, illetőleg nagyobb arányban fordulnak elő, mint a köznyelvi spontán beszédben. Mint láttuk, ennek magyarázata a szociokulturális közegben és bizonyos örökölt nyelvhasználati hagyományokban rejlik. Ezek a sajátosságok is jellemző eszközei a nyelvjárások köznyelvtől való elkülönülésének, s mint ilyenek alkalmasak helyzet- és stílusspecifikálásra, illetőleg társadalmi szimbólumszerepnek a betöltésére, tehát a szűkebb csoporthoz tartozásnak a jelzésére. Tehát „a nyelvjárási szövegeket nemcsak hang- és szóképzetük, hanem szövegössze-függést teremtő eszközeik is jellemzik. Ennek igazolásához [...] főleg szövegek korreferenciahálózatának, forikus elemeinek elemzése szükséges” – írja Dobi Edit (2000: 117). Szükség volna ezen kívül a nyelvjárási szupraszeg-

mentális tartomány számbavételére is, mégpedig a beszélt köznyelvvél való kontrasztjában. Sajnálatos akadályozó tényező azonban, hogy a magyar beszélt köznyelv szóban forgó sajátosságainak a tüzetes leírása mindmáig hiányzik. A magyar nyelvjárások szupraszegmentális sajátosságairól lásd Fodor Katalin munkáit (2001, 2007a és 2007b – továbbvezető irodalmat l. a 2001-es írásban).

Még valamiről röviden! Ez pedig a nyelvjárási szövegek stílusának a kérdése. Tudvalevő, hogy a szövegalkotás és a stílus együtt jár: „A stílus [...] a szövegben nem meghatározottan elkülönülő önálló rész, hanem a szöveg minden pontján tapasztalható nyelvi jelenség” (Tolcsvai Nagy 1996b: 107–108). A nyelvjárási szövegalkotás sajátos stílusa egyebek mellett a fentebb említett jelenségeknek a köznyelvinél nagyobb gyakoriságában van. Abban, hogy elsősorban gyakoriságuk révén váltak és maradtak a köznyelvi nyelvhasználattól megkülönböztető, specifikáló stílusjeggyé. Egy másik stílusforrás a metaforizáció, az erőteljesebb képszerűség nemcsak folklór szövegekben, hanem a hétköznapi beszédben is. Ez különösen a székely és moldvai csángó nyelvjárásokban érhető tetten. Egyetértünk Péntek Jánossal abban is, hogy a szépírói stílus sajátos ízeinek ez az egyik forrása a székely anyanyelvjárású írók esetében Mikestől Tamásiig (l. Péntek 1978, 2000). E témakör kifejtése azonban egy másik tanulmányt igényelne.

Összefoglalás: A nyelvjárási szövegalkotás vizsgálatának a dialektológián túlmutató szerepe

1. A nyelvjárási szövegalkotás vizsgálata a spontán beszélt nyelv kutatásához általában is kitűnő lehetőséget biztosít. A nyelvjárási szövegek stilisztikai kérdései (a stíluskutatók egy részét foglalkoztatták és foglalkoztatják napjainkban is, vö. Bárczi 1961; Sebestyén 1972; Tolcsvai Nagy 1999) az egyik olyan terület, ahol a regionális szövegalkotás kutatása túlmutat a dialektológián. 2. A nyelvjárási szövegfeldvételek a beszélt nyelvi változásvizsgálatra általában jó lehetőséget kínálnak (vö. a Hegedűs-archívummal kapcsolatban megfogalmazott gondolatot: „általa tanulmányozhatók az ötven évvel ezelőtti spontán beszéd jellegzetességei, amelyeket a mai beszéddel összehasonlítva lehetőség nyílik bizonyos változások nyomon követésére” (Nikléczy–Horváth 2007: 177). 3. Hasznosítható eredményeket kínálnak az efféle vizsgálatok a szupraszegmentális terület leírásában is. 4. Általában pedig: „A fonetikai, pszicholingvisztikai vizsgálatok alapjául [...] egyre inkább a spontán beszéd szolgál, amelynek elemzése nemcsak a fonetikai kérdésekre adhat választ, hanem a rejtetten működő beszédtervezési folyamatok pontosabb megismerését is lehetővé teszi” (Nikléczy–Horváth 2007: 173). 5. A nyelvjárási szövegalkotásnak is megvan a maga, kívülről nézve inkább helyzet- és stílus-specifikáló, ezzel pedig a köznyelvtől elkülönítő szerepe. Azt azonban vizsgálatok hiányában nem tudjuk, hogy mekkora ez a szerep. 6. A fontosabb, illetőleg gyakoribb nyelvjárási szövegalkotási jelenségekről – példákat is em-

lítve – főntebb volt szó. 7. A dialektológiában a szövegtani kutatások nálunk mindmáig gyermekcipőben járnak – érthető, de legalábbis magyarázható okokból.

Irodalom

- Andó Éva 2006. A beszélt nyelvi történetmondások elemzésének kognitív és funkcionális szempontjai. In Tolcsvai Nagy Gábor (szerk.): *Szöveg és típus. Szövegtípológiai tanulmányok*. Tinta Könyvkiadó, Budapest. 113–156.
- Balogh Lajos (szerk.) 1993. *Szöveggyűjtemény a táji nyelvváltozatok köréből*. Linguistica Series B. Documenta, 2. MTA Nyelvtudományi Intézet, Budapest.
- Balogh Lajos – Végh József 1975. A Magyar Nyelvjárások Atlaszához kapcsolódó hangfelvételek. In Deme László – Imre Samu (szerk.): *A Magyar Nyelvjárások Atlaszának elméleti-módszertani kérdései*. Akadémiai Kiadó, Budapest. 257–262.
- Bárczi Géza 1961. Nyelvjárás és irodalmi stílus. In *Stilisztikai tanulmányok. A Kiadói Főigazgatóság stilisztikai előadássorozatának teljes anyaga*. Gondolat, Budapest. 62–115.
- Bárth János – Vargha Fruzsina Sára (szerk.) 2011. *Hangok – helyek. Tanulmányok dialektológiai adattárak és helynévtárak számítógépes feldolgozásáról*. ELTE Magyar Nyelvtudományi és Finnugor Intézet, Budapest.
- Büky László 1989–1990. Nyelvjárás és költői nyelvi kitérés. *Nyelv- és Irodalomtudományi Közlemények* 33. 11–14.
- Czine Mihály 1960. *Móricz Zsigmond útja a forradalmakig*. Magvető Kiadó, Budapest.
- DialHandb. 1982–1983. = Besch, Werner – Knoop, Ulrich – Putschke, Wolfgang – Wiegand, Herbert Ernst (Hrsg.): *Dialektologie. Ein Handbuch zur deutschen und allgemeinen Dialektforschung*. Erster und zweiter Halbband. Walter de Gruyter, Berlin–New York.
- Dobi Edit 2000. Nyelvjárás szövegek szövegnyelvészeti vizsgálatához. *Magyar Nyelvjárások* 38. 111–118.
- Fodor Katalin 2001. A magyar nyelvjárások szupraszegmentális sajátosságairól. In Kiss Jenő (szerk.): *Magyar dialektológia*. Osiris Kiadó, Budapest. 346–348.
- Fodor Katalin 2007a. Nyelvjárás intonációs vizsgálatok. In Guttman Miklós – Molnár Zoltán (szerk.): *V. Dialektológiai szimpozion*. A Berzsenyi Dániel Főiskola Magyar Nyelvészeti Tanszékének Kiadványai VIII., Szombathely. 75–79.
- Fodor Katalin 2007b. Nyelvjárás szupraszegmentális sajátosságok. In Zelliger Erzsébet (szerk.): *Nyelv, területiség, társadalom*. A Magyar Nyelvtudományi Társaság Kiadványai 228. Magyar Nyelvtudományi Társaság, Budapest. 213–228.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- Gósy Mária – Gyarmathy Dorottya – Horváth Viktória – Grácsi Tekla Etelka – Beke András – Neuberger Tilda – Nikléczy Péter 2012. BEA: Beszélt nyelvi adatbázis. In Gósy Mária (szerk.): *Beszéd, adatbázis, kutatások*. Akadémiai Kiadó, Budapest. 9–24.
- Hegedűs Attila 2001. A nyelvjárás mondattani jelenségek. In Kiss Jenő (szerk.): *Magyar dialektológia*. Osiris Kiadó, Budapest. 362–373.

- Keszler Borbála 1983. Kötetlen beszélgetések mondat- és szövegtani vizsgálata. In Rácz Endre – Szathmári István (szerk.): *Tanulmányok a mai magyar nyelv szöveg-tana köréből*. Tankönyvkiadó, Budapest. 164–202.
- Kiss Jenő 1982. A szövegszerkesztés néhány jellemző stilisztikai eszköze. In Kiss Jenő (szerk.): *Fejezetek a mihályi nyelvjárás mondattanából*. A Magyar Nyelvtudományi Társaság Kiadványai 164. Magyar Nyelvtudományi Társaság, Budapest. 75–80.
- Kiss Jenő 1993. Észrevételek a nyelvjárási szövegekről. In Fekete Péter – V. Raisz Rózsa (szerk.): *A szöveg szerkesztése, megértése, kidolgozása és megszólaltatása*. A Magyar Nyelvtudományi Társaság Kiadványai 196. Magyar Nyelvtudományi Társaság, Budapest. 50–53.
- Kiss Jenő (szerk.) 2003. *Magyar dialektológia*. Osiris Kiadó, Budapest.
- Kontra Miklós 2013. „A nyelvész eltűnhet” (Walt Wolfram) – és a nyelvi adatok? In Kontra Miklós – Németh Miklós – Sinkovics Balázs (szerk.): *Elmélet és empiria a szociolingvisztikában*. Gondolat Kiadó, Budapest. 253–269.
- Labov, William 1997. Some further steps in narrative analysis. In Bamberg, Michael G. W. (ed.): *Oral versions of personal experience – Three decades of narrative analysis: A special issue of the Journal of Narrative and Life History*. Clark University, Worcester.
- Labov, William – Waletzky, Joshua 1967. Narrative analysis: Oral versions of personal experience. In Helm, June (ed.): *Essays on the verbal and visual arts*. University of Washington Press. 12–44.
- Lanstyák István 2009. *A magyar beszélt nyelv sajátosságai*. Stimul, Pozsony–Bratislava.
- Cs. Nagy Lajos 2013. A számítógép a nyelvjáráskutatásban. *Fórum Társadalomtudományi Szemle* 2013/2. 69–88.
- Nikléczy Péter – Horváth Viktória 2007. Nyelvjárási hangarchívum az interneten. *Beszédkutatás* 2007. 173–178.
- Ochs, Elinor 1979. Planned and unplanned discourse. In Givón, Thomas (ed.): *Syntax and Semantics 12. Discourse and Syntax*. Academic Press, New York. 51–80.
- Ong, Walter J. 2010. *Szóbeliség és írásbeliség. A szó technologizálása*. Alkalmazott Kommunikációtudományi Intézet–Gondolat Kiadó, Budapest.
- Papp Lajos 1983. *Magnetofonnal történő szöveggyűjtés a területi nyelvváltozatok köréből*. Nyelvtudományi Intézet, Budapest. Kézirat, készült Imre Samu és Balogh Lajos tanácsai alapján.
- Péntek János 1978. Rendszerszerűség és produktivitás a Phaseolus népi terminológiájában. *Nyelv- és Irodalomtudományi Közlemények* 22/2.. 175–190.
- Péntek János 2000. *A Magyar dialektológia c. könyv kéziratáról készített lektori vélemény. Kézirat*.
- Radványi Péter 1997–1999. *Multimédiás Tároló Rendszer*. ELTE, Budapest. Helyi sokszorosításban.
- Radványi Péter – Vékás Domokos 1999. *Informatikai lehetőségek a dialektológiában, különös tekintettel a hangzóanyagok kezelésére és a különböző jelegű adatok integrációjára*. ELTE, Budapest. Helyi sokszorosításban.
- Sebestyén Árpád 1972. A tájnyelv mint stílus eszköz Móricz Zsigmond műveiben. *Magyar Nyelvjárások* XVIII. 9–39.

- Silling István 2010. *Vajdasági magyar nyelvjárási olvasókönyv*. Forum Könyvkiadó és Újvidéki Egyetem, Szabadka.
- Simonyi Zsigmond (szerk.) 1902. *Tréfás mesék és adomák. Nyelvjárási olvasókönyv*. Nyelvészeti Füzetek, Athenaeum, Budapest.
- Szabó József 1983. *A mondatszerkesztés nyelvészeti vizsgálata a nagykönyvi nyelvjárási körben*. Akadémiai Kiadó, Budapest.
- Szabó Zoltán 1994. A dialektológia a szövegtan interdiszciplináris keretében. *Nyelv- és Irodalomtudományi Közlemények* 38/1. 55–59.
- Szende Tamás 1973. *Spontán beszédanyag gyakorisági mutatói*. Nyelvtudományi Értekezések 81. Akadémiai Kiadó, Budapest.
- Sziksainé Nagy Irma 1999. *Leíró magyar szövegtan*. Osiris Kiadó, Budapest.
- Tolcsvai Nagy Gábor 1996a. *A nyelvi norma*. Nyelvtudományi Értekezések 144. Akadémiai Kiadó, Budapest.
- Tolcsvai Nagy Gábor 1996b. *A magyar nyelv stilisztikája*. Nemzeti Tankönyvkiadó, Budapest.
- Tolcsvai Nagy Gábor 1999. „Nem találunk szavakat”. *Nyelvértelmezések a mai magyar prózában*. Kalligram, Pozsony.
- Tolcsvai Nagy Gábor 2001. *A magyar nyelv szövegtana*. Nemzeti Tankönyvkiadó, Budapest.
- Tolcsvai Nagy Gábor (szerk.) 2006. *Szöveg és típus. Szövegtipológiai tanulmányok*. Tinta Könyvkiadó, Budapest.
- Vékás Domokos 1999a. *Informatikai lehetőségek a dialektológiában, különös tekintettel a fonetikai szempontokra*. ELTE, Budapest. Helyi sokszorosításban.
- Vékás Domokos 1999b. *Nyelvjárási hanganyagok archiválásáról: egy informatikai alapú koncepció és szellemi kontextusa*. ELTE, Budapest. Helyi sokszorosításban.
- Vincze Krisztina 1986a. Redundanciajelenségek nyelvjárási szövegekben. *Magyar Nyelv* 87. 221–229.
- Vincze Krisztina 1986b. Hiányjelenségek nyelvjárási szövegekben. *Magyar Nyelv* 87. 340–356.

Köszönettel tartozom Balogh Lajosnak szóbeli közléseiért (Balogh Lajos 2014) és a Papp Lajos 1983a. jelzett dokumentumba való betekintés lehetőségéért.

A PALATÁLIS KÖZELÍTŐHANG KÉTFÉLE FUNKCIÓBAN

Gósy Mária

Bevezetés

A szegmentumok, fonémák, betűk mint diszkrét elemek jellemzik a beszédet, és az egymásutániségot sugallják. A folyamatos beszédben azonban a beszédszervek mozgása nem teljesen így szervezett, mivel a kiejtés során az ajkak, a nyelv, a légyszájpad (stb.) folyamatos mozgásban vannak. Ennek eredményeként a beszédhangok artikulációs gesztusai kisebb-nagyobb mértékben átfedik egymást az időben, illetve különféleképpen hatnak egymásra. Ez a jelenség a koartikuláció, amelynek létrejöttét számos tényező befolyásolja, így fiziológiai, fonetikai, fonológiai faktorok, valamint az adott nyelvre jellemző további sajátos mechanizmusok (pl. Beddor et al. 2002). Az utóbbi évtizedek kutatásai azt is igazolták, hogy a beszédhangok egymásra hatása szótaghatáron, sőt szóhatáron túl is érvényesül (pl. Farnetani–Recasens 1999). A koartikuláció több beszédhangon át kiterjedő hatása évtizedek óta ismert és bizonyított (pl. Benguerel–Cowan 1974; Heid–Hawkins 2000). A nyelvek rendkívül változatosak a tekintetben, hogy milyen magánhangzók vagy mássalhangzók követhetik egymást egy szóban (fonotaktikai jellemzők), és hogy ezek az artikulációban hogyan viselkednek, vagyis miképpen és mennyire hatnak egymásra (Öhman 1966; Manuel 1999).

A koartikuláció korai elméletei szerint az egyik beszédhangnak a másik ejtésére gyakorolt hatása megkezdődik, amilyen gyorsan csak lehetséges (pl. Benguerel–Cowan 1974). Más nézetek szerint a jegyterjedés elemzésével írható le a koartikulációs sajátosságok (vö. Hardcastle 2006). Egy további fel fogásban egyfajta időkorlátot feltételeznek a koartikulációs hatással kapcsolatban, ami azt jelenti, hogy egy szegmentum artikulációs gesztusai egy meghatározott szakaszban kezdődnek a szándékozott szegmentum képzését megelőzően (Bell–Berti–Harris 1982). Az ún. akcióelmélet az együttes szerkezetek koncepcióján alapszik, amelyben a beszédprodukciónak egységeit nem szegmentumszintűeknek feltételezik, hanem összetett gesztusoknak (pl. Kelso et al. 1986). Ezek a gesztusok a beszédhangok megformálását szolgálják, és térbeli és időzítési információkat tartalmaznak az artikulációról. Ebben a szemléletben a koartikulációt úgy tekintik, mint a beszédprodukciónak mechanizmus belső mozgásainak automatikus következményét.

Két magánhangzó egymás utáni előfordulásának jelenségét, azaz a köztük lévő ürt a fonetikában és a fonológiában egyaránt hiátusnak nevezik akkor, ha

a szomszédos magánhangzók két szótagmagot alkotnak, és nem jön létre diftongus (Papp 1971; Siptár 2002). Magánhangzók találkozása két szomszédos szótagban a világ számos nyelvében megtalálható, bár akad olyan, amelyik kizárja ennek a lehetőségét, ilyen például a török (Kabak 2007) vagy a kikereve bantu nyelv (Odden 2006). Az angol is megengedi egyes esetekben a két magánhangzó ejtését egymást követő szótagban feloldás nélkül (pl. a *go away* kifejezésben: $CV_1+V_2 > .CV_1.V_2.$). A hiátus kérdése általában és a hiátus megszüntetése mind fonológiai, mind fonetikai szempontból elemezhető; jellemzően több tanulmány található a szakirodalomban a kérdéskör fonológiai (pl. Siptár 2002; Kabak 2007; Padgett 2008), mint fonetikai vetületével (pl. Aguilar 1999; Markó 2012) kapcsolatosan. A hiátus feloldásának módzatai nyelvspecifikusak, amin itt azt értjük, hogy az egyes nyelvekre jellemző, hogy a két magánhangzó kapcsolatát miként és hányféleképpen valósítják meg. A két egymás után következő magánhangzó ejtését a nyelvek többsége igyekszik elkerülni. Ezek az elkerülési módzatok (átmenetek és különféle megszüntetési, feloldási stratégiák) az alábbi hat kategóriába sorolhatók.

- (i) A magánhangzó-kapcsolat feloldása diftongussal ($CV_1+V_2 > .CV_1V_2.$), azaz a két magánhangzó egy szótagot alkot (pl.: *ngbangba* αdz – *.ngba.ngb^a.dz* = 'törpe', a ngiti, Zairéban beszélt, közép-szudáni nyelvben, vö. Casali 1996 vagy *ríos* – *.ríos* = 'folyók' a spanyolban, vö. Aguilar 1999).
- (ii) Az epentézis esete ($CV_1+V_2 > .CV_1.CV_2.$), amikor a két magánhangzó közé egy mássalhangzó ékelődik (pl.: *no-N-pisi-i* – *.nom.pi.sí.ti* = 'söpörni fogok', az axininka perui nyelvben, vö. Casali 1998 vagy *India and* – *ndiərənd* = 'India és' az angolban, vö. Uffmann 2007).
- (iii) A magánhangzó-kiesés esete ($CV_1+V_2 > .CV_2.$ vagy $.CV_1.$), amikor a két magánhangzó egyikét nem ejtik ki (pl.: *deakpa* – *.da.kpa* = 'vásárolj egy bögrét', az etszako, Nigériában beszélt nyelvben, vö. Casali 1998).
- (iv) Siklóhang (glide) létrehozása ($CV_1+V_2 > .CSV_2.$, ahol S a siklóhang) a hiátus feloldására (pl.: *mu-iko* – *.mwii.ko* = 'vakolókanál', az Ugandában beszélt luganda nyelvben). A siklóhang kialakulása két magánhangzó között további szabályok szerint történik számos nyelvben, és jellemzően nem alakul ki ott, ahol a két magánhangzó dallammenete eltérő (Casali 1998).
- (v) Magánhangzók vegyülése ($CV_1+V_2 > .CV_3(:).$) úgy következik be, hogy egy harmadik magánhangzó lesz a két eredetiből, de ez tartalmazza az eredeti magánhangzók egyes artikulációs jegyeit (pl.: *ba.i.ka'atite* – *bee.ka'atite*, a kimatumbi bantu nyelvben, vö. Odden 1996). (Némely nyelvben a vegyülés hosszú magánhangzót eredményez.)
- (vi) Asszimiláció ($CV_1+V_2 > .CV_1.$) *souk* – *.sook* = 'hideg' a törökben, vö. Kabak 2007), amely a magánhangzó megnyúlásával jár együtt. Az asszimilációnak nevezett folyamat a vegyülés eseteként is felfogható, amikor a

kiejtett magánhangzó a két eredeti magánhangzó artikulációs jegyeit tartalmazza, csak különböző mértékben.

Ugyanazon nyelvben a hiátus feloldására különböző mechanizmusok, illetve stratégiák is megjelenhetnek (ilyen például a dél-afrikai xhosa, a kimatumbi bantu nyelv vagy a török, vö. Casali 1996; Odden 1996; Kabak 2007). Éppen ezért a fonológián és a fonetikán kívül a kérdés a grammatika (morfológia) területét is érintheti.

Két szomszédos magánhangzó jelenléte a beszédben már régóta felkeltette a figyelmet. Nem szándékunk ugyan a hiátus feloldását történeti megközelítésben tárgyalni, de megemlíjtük, hogy a hiátustöltőket már évszázadokkal ezelőtt jelölték olykor, hogy tükrözze az ejtést, például a Kulcsár-kódexben (1539, vö. Haader–Papp 1999): *re'am* (306. lap, 6. sor), *reĵam* (40. sor, 20. lap). Mind a felső vessző, mind pedig az *ĵ* betű hangértéke a palatális közelítőhangnak felelt meg. Érdekességgként utalunk Quintilianusra is, a neves rétorra (Kr. után 1. század), aki *Szónoklattanában* (magyarul: 2008) a magánhangzók torlódásáról ír, amelynek bekövetkezéséről azt állapítja meg, hogy ilyenkor „gyengélkedik a beszéd” (625). Javaslatára szerint aggodalom nélkül (!) ejtsük ezeket a magánhangzó-kapcsolatokat.

Két szomszédos magánhangzó kiejtése magyarban is – egyéb nyelvi sajátosságok mellett – függ a magánhangzók minőségétől. A hiátus feloldására elvben három lehetőség adódhat: a magánhangzók egyikének törlése, egy hiátustöltő hang közbeékelése, avagy a két magánhangzó egyikének átalakítása félmagánhangzóvá (Siptár 2002). Közülük a jelen tanulmányban a hiátustöltés egy részjelenségének fonetikai jellemzőivel foglalkozunk. A vonatkozó leírás és kifejtés már Balassa József *Magyar fonétikájában* (1904) megjelenik, aki a két magánhangzó között átmenő „(hang)kapcsokról” ír, amelyeket az angol *glide*-nak (sic) feleltet meg. Leírásában az átmenő kapocs háromféle lehet. Az első a tiszta átmeneti kapocs (ez feltehetően az, amikor mindkét magánhangzót kiejtjük, de nincs közöttük ejtéskönnyítő hang, Balassa jelölésében: $[a^+i]$). A $^+$ jel a két magánhangzó között vélhetőleg egy határozott artikulációs elkülönülést kíván jelezni. A második a hehezetes átmenő kapocs, erre nyelvjárási („népnyelvi”) példát hoz, például *bihal* ('bivaly'). A harmadik a zöngés átmenő kapocs, amelyet a legmegszokottabbként említ, ilyen például az *ai* esete. Hozzáteszi, hogy még szóhatáron is érvényesül, példája: *haza³indul* (a zöngé jelölésére használja a hármas számot; noha a szám feloldását a szerző nem adja meg, a leírás alapján valószínűsíthető, hogy egy glottális zárat szemléltet vele.) A [j]-szerű hang megjelenését inkább népnyelvi ejtésnek gondolja, példája a *házaⁱim* szó (117), ahol a felső indexbe került *i* jésített hangot jelöl. Hozzáteszi, hogy ez utóbbi esetekben az átmenő kapocsból külön kapcsoló hang lett. (A kapcsoló hangot Balassa, úgy tűnik, részben a hangátmenettel azonosítja, l. 123.) A felsoroltakon kívül megengedi a feloldásra a diftongusos ejtést is.

A későbbi szakirodalom többféle terminust használ, és különféle mélységekig részletezi a kérdést. Többen hiátustörvényről, hiátustöltő, illetve ejtés-könnyítő [j] hangról írnak, avagy a jelenséget egyszerűen „hangbetoldásnak” nevezik (utóbbira vö. Elekfi 1992). A mássalhangzó pontos fonetikai jellemzése rendszerint nem található meg (Papp 1971; Kassai 1998). Siptár (2002) „j-szerű mássalhangzó”-nak nevezi a két magánhangzó közé beékelődő hangot. Elekfi a teljes értékűen képzett [j] mássalhangzót mint hiátustöltőt népiesnek ítéli, ejtését a művelt köznyelvben kerülendőnek tartja (1992: 71). Olaszky (2010) a [j]-hez hasonló akusztikai szerkezetű mássalhangzóként jellemzi ezt a hiátustöltőt (136). A hiátustöltő akusztikai sajátosságainak leírását Menyhárt kísérelte meg az időtartam és a formánsszerkezet mérésével (2006). Adatait a szakirodalomban közölt palatális approximánsra megadott értékekhez viszonyította. Eredményei szerint a vizsgált hiátustöltők rövidebbek voltak a nem hiátustöltő mássalhangzóknál, a formánsszerkezetben jelentős különbségeket azonban nem talált. Siptár és Törkenczy elméleti megközelítésben részletezik a hiátustöltők megjelenésének kritériumait, és – többek között – megállapítják, hogy ha a magánhangzó-kapcsolat egyik tagja az [i] vagy az [i:], akkor a beszélők a hiátust kötelezően egy [j]-szerű mássalhangzóval oldják fel. Hozzáteszik, hogy ez független attól, hogy a VV kapcsolat morféma- vagy szóhatáron jelenik-e meg (Siptár–Törkenczy 2000: 283). Különbséget tesznek továbbá a hiátustöltő és a fonémarealizáció minősége között, amennyiben az előbbi gyengébb és inkább átmeneti jellegű az utóbbihoz képest. Ezt a megkülönböztetést főként az ún. szép beszédben vélik jellemzőnek, bár nem tartják kizártnak, hogy ez az eltérés gyors vagy lezser beszédstílusokban alig észrevehető. Elemzéseikben a hiátustöltő és a fonémarealizáció fonetikai különbségét abban határozzák meg, hogy az előbbi siklóhang, az utóbbi pedig likvida. Markó mérések alapján mutatta be a különböző magánhangzó-kapcsolatok ejtési sajátosságait felolvasásban, figyelemmel a megjelenési pozíciókra, valamint elsőként elemezte az irreguláris fonáció megjelenését és feltételezett funkcióját a hiátus eseteiben (2012).

A jelen kutatás azt a kérdést vizsgálta, hogy vajon a hiátustöltők és a /j/ fonéma realizációi között csak funkcióbeli különbség van-e, avagy a funkcióban jelentkező eltérésnek vannak-e ejtési következményei. Amennyiben igen, vagyis az artikulációs gesztusok a hangminőségben is eltéréseket eredményeznek, akkor ez az eltérő ejtés különböző akusztikai következményekkel kell, hogy járjon, azaz a hiátustöltők és a fonémarealizációk nem lesznek azonosak. Ez a kérdés sajátosan összefügg a [j] mássalhangzó osztályozásával is, ami némiképpen bonyolítja a funkció és a hangminőség összefüggésének elemzését a VV kapcsolatokban. A [j]-t hosszú időn keresztül réshangnak tekintették, olykor diftongusnak vélték, pontosabban a kettőshangzó félhangzós elemének (Gósy 2014). Majd approximánsnak, azaz közelítőhangnak (Szende 1992; Gósy 2004), illetve likvidának osztályozták (Siptár 2001). Az akusztikai elemzések alapján hangsúlyozták a magánhangzószerű akusztikai

szerkezetet, és kimutatták az [i] magánhangzó és a [j] akusztikai hasonlóságát. Ez adja az alapot ahhoz a vélekedéshez is, miszerint félmagánhangzó lenne (pl. Szigetvári 1998).

Saját kutatásunk célja a hiátustöltők és a fonémarealizációk között feltételezett beszédhangkülönbség akusztikai-fonetikai elemzése volt azokban az esetekben, amikor a VV kapcsolatok egyik tagját az [i], a másikat pedig vagy az [ɔ], vagy az [ɛ] magánhangzó alkotta. Tekintettel arra, hogy a magyarban ugyanezen magánhangzók között természetesen előfordulhat a palatális approximáns, így azonos környezetben elemezhetővé válik a két mássalhangzó, például *tietek*, illetve *ijesztő*. Arra a kérdésre kerestük a választ, hogy vajon a /j/ fonéma realizációi különböznek-e a hiátustöltőktől a fizikai időtartamukban, valamint az első és a második formánsuk értékében. Hipotézisünk a fonológiai feltételezést követi, vagyis a két mássalhangzó között különbséget feltételeztünk (i) a funkció szerint minden vizsgált paraméterben, (ii) ugyanakkor nem vártunk eltérést attól függően, hogy szó belsejében vagy szóhatáron jelent meg az elemzett mássalhangzó a kétféle funkcióban. Az akusztikai adatok ismeretében állást foglalunk a hiátustöltők és a fonémarealizációk artikulációs gesztusait, valamint hangminőségét illetően.

Kísérleti személyek, anyag, módszer

Tizenhárom magyar anyanyelvű fiatal női beszélő (22 és 32 év közöttiek) vett részt a kísérletben. Valamennyien ép hallásúak, beszédhibájuk nincs, budapestiek, felsőfokú végzettséggel rendelkeznek. A mondatfelolvasások artikulációs tempója 9,8 hang/s és 11,2 hang/s között szórt, az átlag 10,4 hang/s volt.

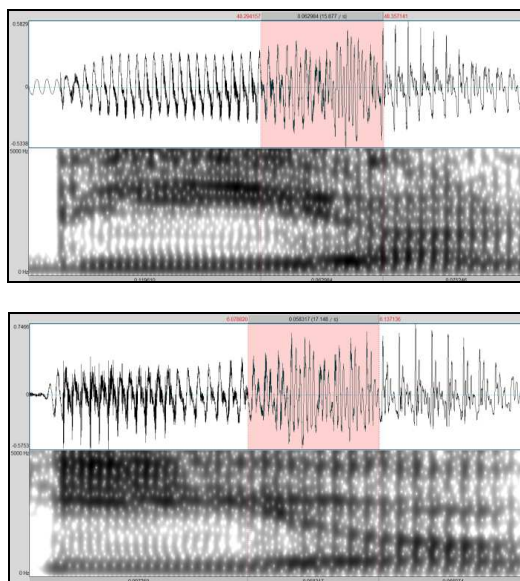
Két és három szótagból álló szavakat, szóösszetételeket és szókapcsolatokat gyűjtöttünk (összesen 80 tételt). A kiválasztott szavak vagy hiátust tartalmaztak (VV), vagy a magánhangzók között a palatális közelítőhangot (V[j]V). A VV minden esetben az [i] és az [ɔ], illetve az [i] és az [ɛ] magánhangzókkal alkotott kapcsolat volt. A V[j]V hangsorokban is ugyanezek a magánhangzók szerepeltek. A lehetséges magánhangzó-kapcsolatok tehát a következők voltak: *ia, ai, ie, ei*, vagyis az [i] hol megelőzte, hol követte a másik magánhangzót. A lehetséges V[j]V hangsorok pedig a következők: *ija, aji, ije, eji*. Egy kategóriát alkottak a szó belsejében megjelenő VV hangkapcsolatok és V[j]V hangsorok (de nem különítettük el a szótőben, illetve a toldalékhatóron megjelenőket). Például *laikus, fiatal, beidéz, tied, Laji, kijavít, helyi, ilyen*. Egy másik kategóriát alkottak a szóösszetételekben és a szóhatáron előfordulókat. Például: *almaital, kocsiaablak, teveitató, pécsi edény, hajillat, kicsi javulás, hely iránt, ovijel*. Ügyeltünk arra, hogy (i) a kiválasztott szavak csak két vagy három szótagúak legyenek, (ii) az elemzendő hangkapcsolatok vagy az első és a második, vagy a második és a harmadik szótag határán forduljanak elő (függetlenül attól, hogy ez szó belsejében vagy szóhatáron realizálódott), (iii) a mindennapokban gyakran előforduló, ismert szavak és szó-

kapcsolatok legyenek. Minden típusban 5 szó fordult elő, egy kategóriában 10 szó, azaz 20 szó szemléltette például az *ei*, *ej* eseteit. A teljes szólista 1040 szót tartalmazott. A szavakat hordozó mondatokba ágyaztuk, amelyek három vagy négy szóból álltak, és a szótagszámuk hasonló volt (6–8). Példák: *Fejlik a tehenet. Piac lesz a közelben. Kefe is kellene. Ilyen filmet nézett. Kicsi javulás látható. Ijesztő képek voltak.*

Noha a V[j]V hangsorokat tartalmazó szavak egy részében az *ly* betű szerepelt (pl. *melyik*); minthogy adatközlőink köznyelvet beszélnek, kijelenthetjük, hogy az írásmód nem járt együtt ejtési eltéréssel. Az elvárásunknak megfelelően a hiátust tartalmazó szavakat a beszélők – a szakirodalomban leírtak alapján – kivétel nélkül feloldották, mégpedig egy palatális közelítőhangra emlékeztető mássalhangzó betoldásával. A /j/ fonéma realizációját tartalmazó szavakban a [j] mássalhangzót ejtették. A véletlenszerű sorrendben leírt 80 mondatot az adatközlők felolvasásában rögzítettük. Az adatközlők feladata az volt, hogy a mondatokat egyetlen levegővel olvassák fel. Bármilyen olvasási hiba esetén az adott mondatot újraolvasták. A hangrögzítés előtt volt idejük a mondatok elolvasására, gyakorlására. Semmilyen egyéb instrukciót vagy magyarázatot a felvétel előtt az adatközlőknek nem adtunk. A mondatokat csendesített szobában digitálisan, közvetlenül számítógépre rögzítettük a GoldWave hangeditáló szoftverrel, 44,1 kHz-es mintavételezéssel. Tárolás: 16 bit, 86 kbit/s, monó. A felvevő mikrofon típusa AT-4040 volt.

A hanganyagot a szerző és egy másik fonetikus elemezte (illetve ellenőrizte). A hangkapcsolatok, valamint a hangsorok fonetikai jellemzői közül a három beszédhang időtartamát, valamint a hiátustöltő és a /j/ realizációinak első és második formánsát mértük meg a Praat 5.3 programban (Boersma–Weenink 2013). Az időtartamok méréséhez szükséges volt a hiátustöltő, illetve a [j] mássalhangzó elhatárolása a megelőző és a követő magánhangzótól. A szakirodalomban használt módszerek eltérők, a magyarra vonatkozóan Olaszky (2006; 2010) a hallás alapján történő szegmentálást alkalmazta. Az [i] és [j] beszédhangok elkülönítése a vizuális leképeződés (hangszínkép és oscillogram) alapján számos nehézséget vet fel (l. Vértes O. 1982), ezért igyekeztünk igen körültekintően eljárni, és minden egyes mérés esetében szigorúan ugyanazt az eljárást és beállítást alkalmaztuk. Anyagunkban a hiátustöltők és a /j/-realizációk határainak kijelöléséhez a Praat következő beállításait használtuk. Frekvenciatartomány: 0–5000 Hz, ablakszélesség (window length): 0,005 s, window shape: Gaussian, dinamikatartomány (dynamic range): 50 dB; az alaphangmagasság tartománya (pitch range): 70–350 Hz, a frekvenciasávok száma 513, frekvencialépés szélessége: 43,1 Hz. A szegmentálás fő kritériumát a második formáns menetének megváltozásában határoztuk meg (mind a Praat automatikus kiíratását, mind az elemző saját vizuális megítélését tekintetbe véve), figyelembe vettük továbbá a rezgéskép információját, ha szükséges volt, az első formáns változását. A magánhangzók kezdetének és végének megállapítását a megelőző, illetve követő mással-

hangzók akusztikai szerkezete is segítette; itt szokásosan a második formánsok első, illetve utolsó rezgését tekintettük irányadónak. A pontosabb elhatárolás szükségessége esetén az intenzitás és az alaphangmagasság alakulását is figyelembe vettük. Minden esetben hallás alapján is ellenőriztük a kijelölt hangokat. Az 1. és a 2. ábra a *dia/kija* hangkapcsolat, illetve hangsor ejtéséről készült regisztrátumokat szemlélteti szóban és szóhatáron.



1. ábra

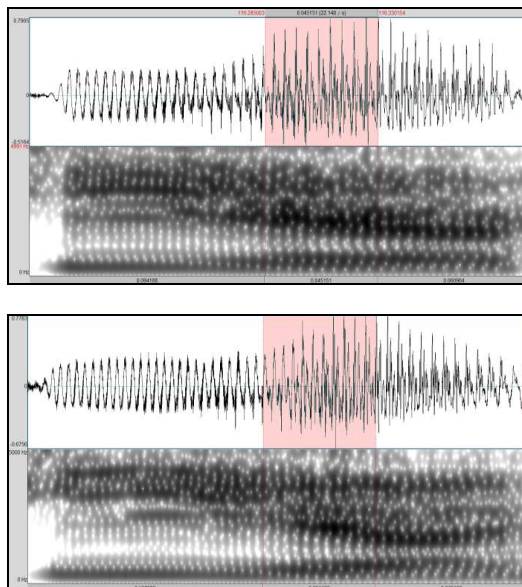
A *dia* szó ejtéséről (felső kép) és a *kija* hangrészlet ejtéséről (a *kijavít* szóból, alsó kép) készült rezgékép és hangszínkép (a szürke sáv a hiátustöltőt, ill. a közelítőhangot jelöli)

Az ábrák szemléltetik a hiátustöltőt, illetve a palatális közelítőhang időtartamát az *ia*, illetve *ija* eseteiben szón belül és szóhatáron (ugyanazon adatközlőktől mindkét ábrán). A formánsokat a hiátustöltő és a /j/-realizációk megállapított időtartamának közepén mértük, amit FFT-elemzéssel egészítettünk ki. A statisztikai vizsgálatokhoz ismétléses ANOVA-t, valamint korrelációelemzést használtunk az SPSS (19.0) programban, 95%-os konfidenciaszinten.

Eredmények

Először az időtartamadatokat, majd a formánsértékek alakulását mutatjuk be. A statisztikai elemzésekben a hiátustöltő és a fonémarealizáció különbsé-

gét a 'funkció' tényezőjeként jelenítjük meg. Mind az időtartam, mind az első és a második formánsok értékeinek elemzésekor a 'funkció' mellett vizsgáltuk a 'hangkapcsolat' és a 'szóhatár' tényezőket is. A 'hangkapcsolat' mint tényező arra utal, hogy vajon az alsó nyelvállású palatális, illetve veláris magánhangzók minősége, valamint az [i] sorrendisége hatással van-e a vizsgált paraméterek értékeinek alakulására. A 'szóhatár' mint tényező pedig arra keres választ, hogy az elemzett hangsorok pozíciója – a tekintetben, hogy szó belsejében vagy szóhatáron jelennek-e meg – befolyásolja-e az elemzett paraméterek alakulását.



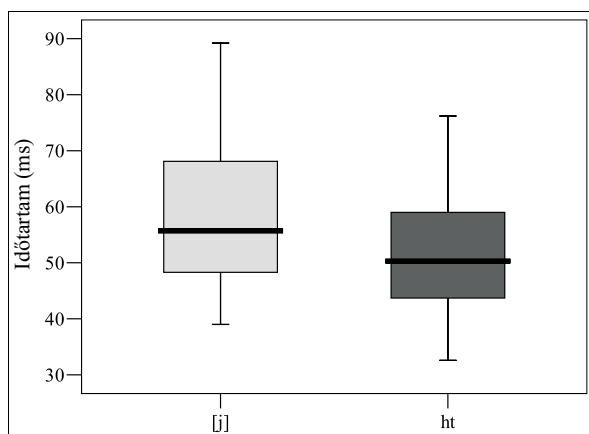
2. ábra

A *kia* részlet ejtéséről (a *Ki a rokonod itt?* mondatból) (felső kép) és a *kija* hangrészlet ejtéséről (a *Ki Janics Natasa?* mondatból, alsó kép) készült rezgéskép és hangszínekép (a szürke sáv a hiátustöltőt, ill. a közelítőhangot jelöli)

Időtartamok

A kiinduló kérdés a hiátustöltő és a palatális közelítőhang fizikai időtartamának értékeit érintette. Adataink azt mutatták, hogy – minden egyéb tényező figyelembevétele nélkül – a fizikai időtartamok szignifikánsan különböznek a 'funkció' függvényében [$F(1, 104) = 69,417$; $p < 0,001$; $\eta^2 = 0,42$]. Ez a tényező relatíve nagymértékben magyarázza az időtartamok különbségét (mint láttuk, 42%-ban). A hiátustöltők átlagos időtartama 52 ms (átl. elt.:

10,43 ms), míg a /j/-realizációké 58,5 ms (átl. elt.: 11,27 ms). A minimumérték a hiátustöltők esetében 32,6 ms, a /j/-realizációké 39 ms, a maximumérték az előzőeknél 76,2 ms, az utóbbiaknál 89,2 ms (3. ábra).



3. ábra

A hiátustöltők (ht) és a /j/-realizációk időtartamai (medián és szórás)

Részletesebben elemeztük a hangkapcsolat, illetve a hangsor fonetikai jellemzőinek esetleges hatását; azt, hogy a magánhangzó minősége, illetve az [i] helye (az, hogy megelőzi vagy követi a másik magánhangzót) befolyásolja-e a fizikai időtartamok alakulását. Az 1. táblázat az adatokat tartalmazza a VV és a V[j]V típusainak függvényében. A 'hangkapcsolat' mint tényező nincs szignifikáns hatással az időtartamok alakulására.

Az átlagértékek nagyon hasonlóak a hangkapcsolat és a hangsor valamennyi formájában. Mind a hiátustöltő, mind a /j/-realizáció valamivel hosszabb időtartamú akkor, ha az [i] az első hang a VV kapcsolatban, illetve ha megelőzi a /j/-t. A különbségek azonban csupán 1–4 ms közöttiek. Az [i] helyzetétől függően a vizsgált [j]-k minimumértékei hosszabbak az *aji* és *eji*, mint az *ija* és *iye* esetében. Az *ai* és az *ia* hangkapcsolatoknál azonban nincsen ilyen eltérés; az *ei* pedig rövidebb, mint az *ie*. A különbségek itt is nagyon kicsik, a fonémarealizációknál nem éri el a 4 ms-ot, az *ie* esetében 4,6 ms. A maximumértékek tendenciája hasonló; a különbség az *aji/ija* hangsorban valamivel nagyobb (7,4 ms), mint a másik magánhangzó jelenléte esetében (2,8 ms), az *ai/ia* különbsége 6 ms. Az *ei* értékei hasonlóan alakulnak, mint az *ie* esetében, de a különbség mindössze 1,6 ms.

A 'szóhatár' mint tényező matematikailag igazolható hatással van a hiátustöltők, illetve a /j/-realizációk időtartamára [$F(1, 104) = 4,653$; $p = 0,033$; $\eta^2 = 0,046$]. A tényező azonban csak igen kismértékben magyarázza az elté-

réseket (mértéke 5% alatti). A hiátustöltő időtartamának átlaga szóban 53,5 ms (átl. elt. 10,64 ms), szóhatáron 50,6 ms (átl. elt. 10,11 ms); a /j/-realizációké szóban 58,3 ms (átl. elt. 11,88 ms), szóhatáron 58,6 ms (átl. elt. 10,75 ms). A magánhangzók minőségétől és sorrendjétől függetlenül a /j/ realizációi hosszabb időtartamban valósulnak meg szóhatáron a szó belseji pozícióhoz képest, a hiátustöltők ugyanakkor rövidebb időtartamúak szóhatáron, mint szó belsejében (2. táblázat). Az átlagok különbsége a /j/-realizációknál 0,3 ms, vagyis gyakorlatilag nincs különbség a szóban elfoglalt helyzet és a szóhatáron előforduló ejtések időtartama között. A hiátustöltők időtartamainak átlagai közötti eltérés is nagyon kicsi, 2,9 ms, átlagosan ennyivel ejtik hosszabban az adatközlők a hiátustöltőket a szó belsejében (4. ábra).

1. táblázat: A fonémarealizációk és a hiátustöltők időtartama a magánhangzós kontextus függvényében

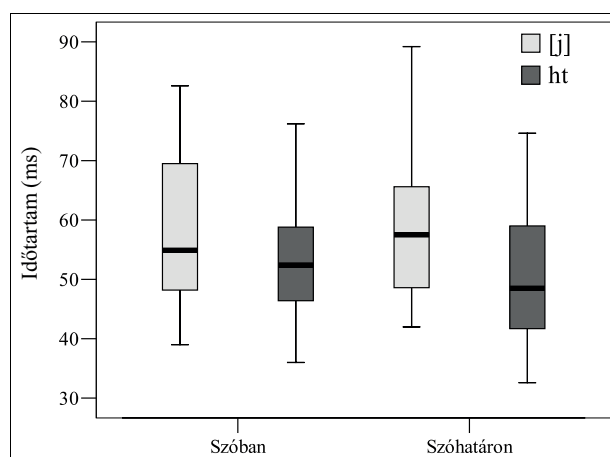
Magánhangzós kontextus	Időadatok megnevezése	/j/-realizációk (ms)	Hiátustöltő (ms)
<i>ai</i>	átlag	58,90	51,50
	átlagos eltérés	11,59	10,12
	minimumérték	43,60	36,20
	maximumérték	81,80	69,20
<i>ia</i>	átlag	60,20	54,80
	átlagos eltérés	12,65	10,86
	minimumérték	39,80	36,00
	maximumérték	89,20	75,20
<i>ei</i>	átlag	56,20	49,90
	átlagos eltérés	9,10	11,24
	minimumérték	42,20	32,60
	maximumérték	75,80	76,20
<i>ie</i>	átlag	58,60	52,00
	átlagos eltérés	11,71	9,41
	minimumérték	39,00	37,20
	maximumérték	78,60	74,60

A szó belsejében és a szóhatáron megjelenő mássalhangzók időadatait részleteztük a részt vevő magánhangzók és azok sorrendje szerint. A statisztikai eredmény szignifikáns eltérést mutatott [$F(3, 104) = 4,734$; $p = 0,004$; $\eta^2 = 0,129$], ami azt jelenti, hogy a hiátustöltő, illetve a /j/ fonéma realizációinak időtartamát mind a hangkapcsolat típusa, mind pedig a pozíciója meghatározza. A [j] kétféle funkcióját tekintve azonban sem a hangkapcsolat, sem a szóhatár nem igazolt szignifikáns eltérést, ami azt jelenti, hogy a hangkapcsolat típusától és a pozíciótól függetlenül a hiátustöltő és a /j/-realizációk között nem volt matematikailag alátámasztott időbeli különbség. A 3. táblázat részletes adatai szerint mind a hiátustöltők, mind a /j/-realizá-

ciók kismértékben hosszabbak szó belsejében a szóhatáron megjelenőkhöz viszonyítva. Ez azonban csak az *ai*, *aji* és *ia*, *ija* eseteire igaz. Az *eji* és az *ije* eseteiben a fonémarealizációk éppen ellenkező tendenciát mutatnak, itt tehát a szó belsejében hosszabb a mássalhangzó. A hiátustöltő a hangkapcsolat függvényében is eltér; az *ei* esetében szóban hosszabb a mássalhangzó, az *ie* kapcsolatban pedig szóhatáron. A különbségek mindenütt igen kicsik, néhány ms-osak.

2. táblázat: A fonémarealizációk és a hiátustöltők időtartama a pozíciótól függően

Pozíció	Időadatok megnevezése	/j/-realizációk (ms)	Hiátustöltő (ms)
Szó belsejében	átlag	58,30	53,50
	átlagos eltérés	11,88	10,64
	minimumérték	39,00	36,00
	maximumérték	82,60	76,20
Szóhatáron	átlag	58,66	50,60
	átlagos eltérés	10,75	10,11
	minimumérték	42,00	32,60
	maximumérték	89,20	74,60



4. ábra

A hiátustöltők (ht) és a /j/-realizációk ([j]) időtartamai szó belsejében és szóhatáron (medián és szórás)

A hiátustöltők és a fonémarealizációk temporális viszonyai jellemzők lehetnek a tekintetben is, hogy azok időtartama hogyan viszonyul a teljes VCV

hangsor időtartamához. Adataink azt mutatták, hogy nincs statisztikailag igazolható különbség az időtartamarányokban a kétféle funkció között. A /j/-realizációk átlagos aránya a V[j]V hangsor teljes idejéhez képest 30,2% (átl. elt. 3,98%), a hiátustöltők pedig 30,6% (átl. elt. 3,99%). A minimum- és a maximumértékek aránya is nagyon hasonló a kétféle funkcióban: a /j/ megvalósulásaiban 22,1%, illetve 41,4%, a hiátustöltők esetében pedig 21,6%, illetve 38,9%. A 'funkció' és a 'hangkapcsolat', valamint a 'funkció' és a 'szóhatár' tényezők sincsenek matematikailag igazolható hatással az időtartamarányokra. Elemeztük a 'funkció', a 'hangkapcsolat' és a 'szóhatár' tényezők együttes hatását, és itt sem kaptunk szignifikáns különbséget. A mássalhangzók időtartamarányai a teljes VCV hangsor időtartamához képest tehát a funkciótól és minden egyébtől függetlenül azonosak, illetve statisztikailag nem különböznek.

3. táblázat: A fonémarealizációk és a hiátustöltők időtartama a magánhangzós környezet és a pozíció függvényében

Funkció	Hangkapcsolat típusa	Pozíció	Időtartam átlaga (ms)	Időtartam átlagos eltérése (ms)
/j/ realizációi	ai	szóban	59,60	12,25
		szóhatáron	58,10	11,34
	ia	szóban	63,10	12,50
		szóhatáron	57,40	12,63
	ei	szóban	53,40	9,38
		szóhatáron	58,90	8,27
	ie	szóban	57,00	12,28
		szóhatáron	60,30	11,38
	ai	szóban	52,90	9,87
		szóhatáron	50,10	10,58
hiátustöltők	ia	szóban	56,40	11,82
		szóhatáron	53,10	10,02
	ei	szóban	53,70	11,79
		szóhatáron	46,10	9,67
	ie	szóban	51,10	9,45
		szóhatáron	52,90	9,68

Formánsszerkezet

Mint említettük, a statisztikai vizsgálatban a 'funkció' a hiátustöltő, illetve a /j/-realizáció különbségét, a 'hangkapcsolat' a VV, illetve V[j]V magánhangzóinak típusait, a 'szóhatár' tényező pedig azt jelentette, hogy a vizsgált mássalhangzó szó belsejében vagy két szó határán valósult-e meg. A 'funkció' tényező nincs hatással az **első formánsok** értékére (szignifikáns különbség nem volt kimutatható közöttük). Az összes /j/-realizáció első formánsai-

nak átlaga 525 Hz (átl. elt. 48,39 Hz), a hiátustöltőké 529 Hz (átl. elt. 45,47 Hz). Az előbbieknél a minimumérték 386 Hz, a maximumérték pedig 621 Hz, az utóbbiaknál az első formánsok minimuma 407 Hz, a maximuma pedig 626 Hz volt.

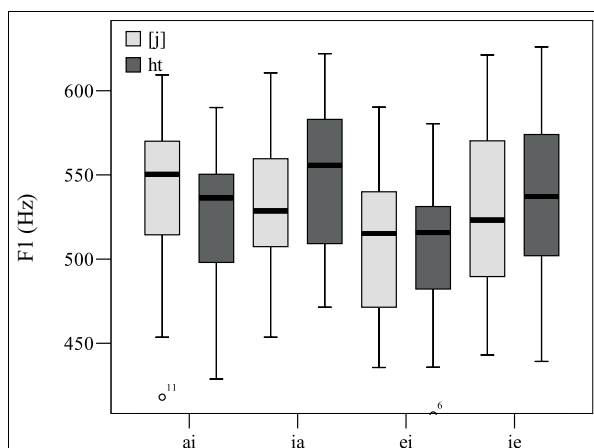
A 'hangkapcsolat' és a 'funkció' együttesen szignifikáns hatással van az F_1 értékeire [$F(1, 104) = 5,841$; $p = 0,001$; $\eta^2 = 0,154$]. Ez azt jelenti, hogy mind az [i] elhelyezkedése, mind a másik két magánhangzó minősége befolyásolja a /j/-realizációk és a hiátustöltők első formánsainak alakulását (4. táblázat).

4. táblázat: A fonémarealizációk és a hiátustöltők első formánsainak adatai a magánhangzók minőségének és sorrendjének függvényében

VV és V[j]V típusai	F_1 adatainak megnevezése	/j/-realizációk F_1 -értéke (Hz)	Hiátustöltő F_1 -értéke (Hz)
ai	átlag	537,20	523,70
	átlagos eltérés	47,13	40,50
	minimum	418,00	428,80
	maximum	609,40	590,00
ia	átlag	527,00	545,90
	átlagos eltérés	49,56	42,80
	minimum	386,00	471,40
	maximum	610,60	622,00
ei	átlag	507,80	505,40
	átlagos eltérés	42,55	38,77
	minimum	435,60	407,40
	maximum	590,20	580,40
ie	átlag	527,90	539,40
	átlagos eltérés	51,77	50,24
	minimum	443,00	439,20
	maximum	621,20	626,00

A részletes adatok azt mutatják, hogy a fonémarealizációk esetén az első formánsok csaknem azonosak az *aji/ija* eseteiben, az utóbbi átlagosan 10 Hz-cel alacsonyabb. Az *ejí/ije* ejtésekor valamivel nagyobb a különbség, de itt az *ije* ejtésekor tapasztaltuk a magasabb értéket, a különbség azonban itt sem nagy, átlagosan 20 Hz. A hiátustöltőknél a tendencia az *ai/ia* esetében ellentétesen alakult; az *ia* kapcsolatban átlagosan magasabb volt az első formáns értéke, a különbség 22,2 Hz. Az *ei/ie* ejtései hasonlóak az *ejí/ije* realizációihoz, a hiátustöltők eseteiben is az utóbbinál volt magasabb F_1 -érték, a különbség 34 Hz (5. ábra). Általánosságban megállapíthatjuk, hogy az elemzett mássalhangzó első formánsai magasabbak, ha az [i] megelőzi a másik magánhangzót. A hiátustöltők ejtésében az első formáns értékei azonban állandóbb tendenciát mutatnak, függetlenül a magánhangzók minőségétől, illetve az [i] helyzetétől. Az ugyanazon hangsoron belüli szórás valamivel

nagyobb mértékű, mint az eltérő hangsorok közötti frekvenciakülönbség (5. táblázat).



5. ábra

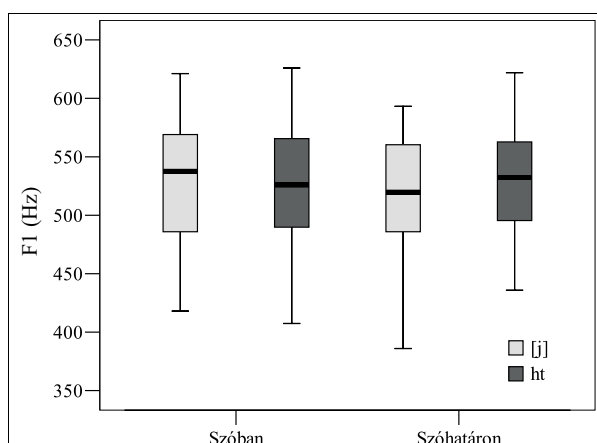
A fonémarealizációk ([j]) és a hiátustöltők (ht) első formánsainak frekvenciaértékei (medián és szórás)

5. táblázat: A fonémarealizációk és a hiátustöltők első formánsainak értékei szó belsejében és szóhatáron

Pozíció	F ₁ adatainak megnevezése	/j/-realizáció F ₁ -értéke (Hz)	Hiátustöltő F ₁ -értéke (Hz)
Szó belsejében	átlag	531,80	525,60
	átlagos eltérés	48,75	45,99
	minimumérték	418,00	407,40
	maximumérték	621,20	626,00
Szóhatáron	átlag	518,10	531,70
	átlagos eltérés	47,50	45,18
	minimumérték	386,00	435,80
	maximumérték	593,20	622,00

A 'funkció' és a 'szóhatár' tényezők együttesen ugyancsak meghatározók a vizsgált mássalhangzók első formánsainak alakulására, a hiátustöltők és a fonémarealizációk közötti különbség szignifikáns [$F(1, 104) = 10,937$; $p = 0,001$; $\eta^2 = 0,102$]. A fonémarealizációk és a hiátustöltők átlagértékeinek különbsége szó belsejében 6,2 Hz; szóhatáron valamivel nagyobb, 13,6 Hz. A /j/ megvalósulásaiban az első formáns átlagának különbsége a pozíciótól függően 13,7 Hz, a hiátustöltők esetében ez a különbség átlagosan 6,1 Hz. A funkciótól függően

tehát az értékek kismértékben változnak; a fonémarealizációk F_1 -e valamivel magasabb szó belsejében, míg a hiátustöltőké szóhatáron kissé magasabb, a különbségek azonban mindkét funkcióban nagyon kicsik (6. ábra).



6. ábra

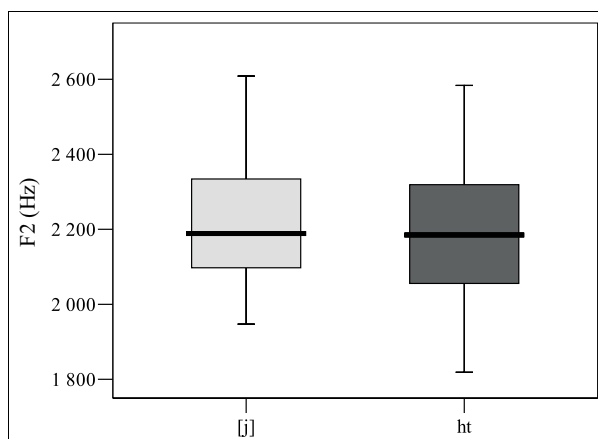
Az első formáns értékei szó belsejében és szóhatáron a funkciótól ([j], illetve hiátustöltő „ht”) függően (medián és szórás)

6. táblázat: A fonémarealizáció és a hiátustöltő első formánsainak értékei a hangkapcsolat típusának és a pozíciónak a függvényében

Funkció	Hangkapcsolat	Pozíció	F_1 átlaga (Hz)	F_1 átlagos eltérés (Hz)
/j/-realizáció	<i>aji</i>	Szóban	531,30	52,86
		Szóhatáron	543,20	41,90
	<i>ija</i>	Szóban	536,60	47,37
		Szóhatáron	517,50	51,73
	<i>ejj</i>	Szóban	512,90	40,21
		Szóhatáron	502,60	45,79
hiátustöltő	<i>ije</i>	Szóban	546,50	52,89
		Szóhatáron	509,20	45,08
	<i>ai</i>	Szóban	533,60	46,15
		Szóhatáron	513,80	32,78
	<i>ia</i>	Szóban	534,40	41,70
		Szóhatáron	557,50	42,30
	<i>ei</i>	Szóban	500,50	40,54
		Szóhatáron	510,40	37,88
	<i>ie</i>	Szóban	533,90	50,67
		Szóhatáron	544,90	51,25

A statisztikai elemzések szignifikáns különbséget igazoltak az első formáns értékére az összes vizsgált tényező – 'funkció', 'hangkapcsolat', 'szóhatár' – figyelembevételkor [$F(3, 104) = 9,290$; $p = 0,001$, $\eta^2 = 0,225$]. Ennek a vizsgálatnak a részletes adatait a 6. táblázat összegzi. A funkciók szerinti csoportokban (/j/-realizáció, ill. hiátustöltő) azonban sem a hangkapcsolatnak, sem a szóhatárnak nem volt szignifikáns hatása az első formánsok értékére.

A **második formánsok** statisztikai elemzése szignifikáns különbséget mutatott a 'funkció' tényezőre, vagyis matematikailag igazolható a fonémarealizációk és a hiátustöltők F_2 -értékeinek különbsége [$F(1, 104) = 10,841$; $p = 0,001$; $\eta^2 = 0,101$]. A /j/-k összes megvalósulásaiban az F_2 -k átlaga 2218 Hz (átl. elt. 158,84 Hz), a hiátustöltőké 2189 Hz (átl. elt. 174,71 Hz). Az előbbieket minimum- és maximumértékei 1947 Hz és 2609 Hz, az utóbbiaké 1819 Hz és 2584 Hz. Az átlagértékek között 39 Hz-nyi az eltérés; a fonémarealizációk második formánsai kismértékben magasabb értékeken realizálódnak (7. ábra).



7. ábra

A fonémarealizációk ([j]) és a hiátustöltők (ht) második formánsainak frekvenciaértékei (medián és szórás)

A 'funkció' és a 'hangkapcsolat' tényezők együttesen nincsenek hatással az adatokra, a formánsértékek különbségei elenyészők. A /j/-realizációk második formánsainak átlagai az *aji/ija* hangsorokban 43 Hz különbséget mutatnak, az *aji/ije* eseteiben pedig 36 Hz-nyit. A hiátustöltők második formánsai a hangkapcsolatok függvényében mindössze 4 Hz-ben térnek el átlagosan. Valamennyi F_2 -értéknél a szórás is nagyon hasonló tartományt mutat, függetlenül a magánhangzós környezettől. A minimum- és a maximumértékek

ugyanakkor egyes hangsorokban jellegzetesen különböznek. Az *aj*-ban a mássalhangzó F_2 -maximuma 90 Hz-cel magasabb, mint az *ija* mássalhangzó-jáé, az *ej*-ben viszont ugyancsak 90 Hz-cel, de alacsonyabb, mint az *ije* mássalhangzójának második formánsa. A hiátustöltők F_2 -inek minimumértéke az *ai* esetben szintén közel 90 Hz-cel magasabb az *ia* hangkapcsolathoz képest, ugyanakkor a maximumérték ez utóbbiban 83 Hz-cel haladja meg az előző maximumát. Az *ei* hiátustöltőinek minimumértéke 112 Hz-cel magasabb, mint az *ie* hiátustöltőié, a maximumértékek 143 Hz-es különbséget mutatnak (az *ie* esetben jelentkezett a magasabb érték).

Az F_2 frekvenciatartománya tendenciaszerűen kissé eltér a funkció függvényében. A fonémarealizációknál a minimumértékek azonosnak tekinthetők; a maximumértékek különbsége megközelíti, de nem éri el a 100 Hz-et. A hiátustöltőknél mind a minimum-, mind a maximumértékek nagyobb mértékben térnek el a hangkapcsolat típusának (magánhangzó minősége és sorrendjük) függvényében. A minimum F_2 -k magasabbak az *ai* és az *ei* esetekben, a maximumok pedig az *ia* és az *ie* hangkapcsolatokban, itt tehát egyfajta „kiegyenlítődség” történik. Ha összehasonlítjuk a maximumértékeket a kétféle funkcióban, akkor azt látjuk, hogy az *ej/ije* és az *ei/ie* hangsorokban a tendencia azonos. Az *aji/ai* és az *ija/ia* hangsorokban ezek az értékek ugyanakkor éppen ellentétes tendenciát mutatnak.

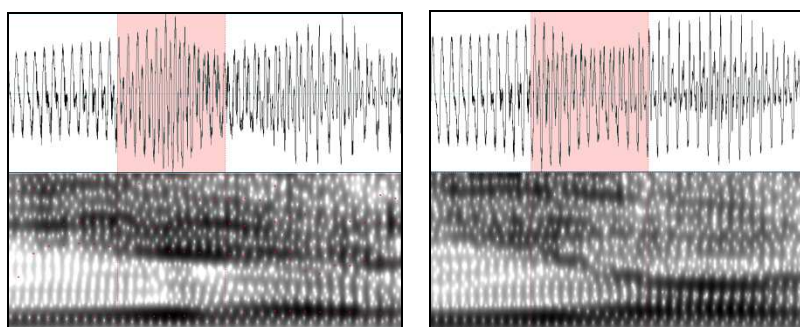
Elemeztük, hogy a palatális és a veláris alsó nyelvvállású magánhangzók esetleges hatása kimutatható-e a mássalhangzó F_2 -értékeiben (8. táblázat). A magánhangzók sorrendjétől függetlenül a különbség a /j/-realizációk átlagában 211 Hz, a hiátustöltők esetében pedig átlagosan 237 Hz (az átlagos eltérés értékei nagyon hasonlóak). Mindez arra utal, hogy a magánhangzó minősége hatással látszik lenni mind a fonémarealizációk, mind a hiátustöltők második formánsára. A palatális magánhangzó környezetében magasabbak az értékek, mint a veláris esetében (a különbségek azonban statisztikailag nem voltak igazolhatók, feltehetően a nagy szórás miatt).

A második formánsok alakulását a magánhangzók hatására a 8. ábra szemlélteti. A két hiátustöltő F_2 -értékének különbségét a szürke kiemelés mutatja.

A 'szóhatár' mint faktor szignifikáns eltérést igazolt az F_2 frekvenciaértékére [$F(1, 104) = 14,989$; $p = 0,001$; $\eta^2 = 0,135$], ami azt jelenti, hogy a szó belsejében, illetve a szóhatáron előforduló hangsorokban a második formánsok értéke matematikailag igazolhatóan különböző volt. A /j/-realizációk F_2 -inek átlagos különbsége mindössze 22 Hz-nek adódott, a minimálisan magasabb érték a szóhatáron előfordulóakra volt jellemző. A hiátustöltők második formánsainak átlaga ugyanakkor a szóhatáron kissé alacsonyabb, mint szó belsejében, a különbség 49 Hz (9. táblázat). A tendencia tehát a funkciótól függően ellentétes. A fonémarealizációk és a hiátustöltők második formánsainak átlaga szó belsejében csaknem azonos (a különbség 6 Hz). Szóhatáron a fonémarealizációk F_2 -inek átlaga magasabb, mint a hiátustöltőké, a különbség 65 Hz (9. ábra).

8. táblázat: A /j/-realizációk és a hiátustöltők második formánsainak adatai a magánhangzós környezet függvényében

Funkció	V helyzete az [i]-hez képest	V	Átlag (Hz)	Átlagos eltérés (Hz)
/j/-realizáció	megelőzi	[ɔ]	2091	109,11
		[ɛ]	2341	102,50
	követi	[ɔ]	2134	102,58
		[ɛ]	2306	152,36
	összes	[ɔ]	2113	107,06
		[ɛ]	2324	129,80
hiátustöltő	megelőzi	[ɔ]	2073	100,80
		[ɛ]	2308	94,96
	követi	[ɔ]	2068	156,23
		[ɛ]	2305	155,61
	összes	[ɔ]	2070	130,19
		[ɛ]	2307	127,65



8. ábra

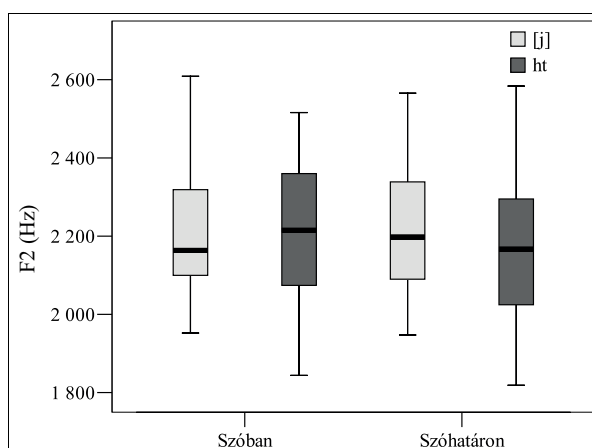
A *kies* (bal oldalon) és a *miatt* (jobb oldalon) szavak ejtéséről készült rezgéskép és hangszínekép (ugyanazon beszélő ejtésében)

A második formánsok esetében a 'funkció', a 'hangkapcsolat' és a 'szóhatár' tényezők együttesen nem gyakorolnak szignifikáns hatást a hangsorokban megjelenő mássalhangzók második formánsára. A funkciókülönbség szerinti egyes csoportokon belül sem a 'hangkapcsolat', sem a 'szóhatár' faktorok nem szignifikánsak. Ennek ellenére röviden jellemezzük a kapott adatokat a magánhangzó-környezet függvényében. Mind a fonémarealizációk, mind a hiátustöltők második formánsainak átlagértékei a szó belsejében és a szóhatáron megjelenő mássalhangzóban 100 Hz-nél kisebb mértékben különböznek valamennyi hangkapcsolattípus esetében. Egyáltalán nem találtunk

eltérést az átlagokban például az *aji/ija* hangsorokban és az *ai/ie* hangkapcsolatokban sem a szó belseji, sem a szóhatáron történő előfordulásban; illetve 50-60 Hz-nyi átlagos különbséget adatoltunk az *ije/ie* eseteiben. A legnagyobb különbség (78 Hz) az *ia* hiátustöltőjének F₂-jében adódott a szóhatár függvényében, a szó belsejében megjelenő mássalhangzóban volt magasabb az érték.

9. táblázat: A második formánsok adatai a mássalhangzó funkciójának és pozíciójának függvényében

Pozíció	F ₂ adatainak megnevezése	/j/-realizációk F ₂ -értékei (Hz)	Hiátustöltők F ₂ -értékei (Hz)
szó belsejében	átlag	2207,00	2213,00
	átlagos eltérés	148,03	172,40
	minimumérték	1952,00	1844,00
	maximumérték	2609,00	2516,00
szóhatáron	átlag	2229,00	2164,00
	átlagos eltérés	169,76	175,23
	minimumérték	1947,00	1819,00
	maximumérték	2566,00	2584,00



9. ábra

A fonémarealizációk ([j]) és a hiátustöltők (ht) második formánsainak értékei (medián és szórás)

Valamennyi adat figyelembevételével választ kerestünk arra a kérdésre, hogy vajon a **magánhangzós kontextus** melyik összetevője (azaz a V-k minősége, illetve az [i] helyzete) gyakorol-e nagyobb hatást a /j/-realizációk és a

hiátustöltők időtartamaira és formánsaira. A statisztikai eredmények azt mutatták, hogy a magánhangzós kontextusnak egyik funkcióban sincs hatása sem az időtartamra, sem a második formánsra. Az időtartam a magánhangzós környezet függvényében mindkét funkcióban néhány milliszekundumos eltérést mutat. A második formánsok kissé magasabbak a palatális alsó nyelvéllátású magánhangzó környezetében, mint a veláris alsó nyelvéllátású környezetében. Az első formáns alakulását ugyanakkor szignifikánsan befolyásolja az [i] pozíciója [$F(1, 104) = 11,134; p = 0,001; \eta^2 = 0,10$]. Ez utóbbi különbségei a fonémarealizációknál 10-20 Hz-esek, a hiátustöltőknél pedig 10 Hz alattiak.

Az elemzett mássalhangzó idő-tartamának **összefüggéseit** a megelőző és a követő magánhangzó időtartamával, valamint az idő-tartamoknak és a formánsoknak az összefüggéseit korrelációelemzéssel ellenőriztük. A /j/ megvalósulásainak időtartama és a megelőző magánhangzó időtartama között gyenge közepes korreláció volt kimutatható ($r = 0,326; p = 0,001$). A hiátustöltő időtartama a megelőző magánhangzó időtartamával közepesen erős ($r = 0,471; p = 0,001$), a követő magánhangzó időtartamával gyenge korrelációt mutatott ($r = 0,228; p = 0,05$). A szomszédos magánhangzók hatása nem túlzottan erős, de mindkét funkcióban meghatározó a megelőző magánhangzó hatása, vagyis a progresszív hatás jobban érvényesül. A hiátustöltők esetében kismértékben tekintetbe vehető a követő magánhangzó hatása is. A mássalhangzó formánsai és a magánhangzók időtartamai között nem találtunk összefüggést.

Következtetések

Az eredmények alapvetően alátámasztották a hipotézist, amennyiben az időtartamok és a formánsok számos tekintetben statisztikailag különböztek a funkció függvényében. A kapott értékek azonban csekély eltéréseket igazoltak; az időtartamoknál kevesebb mint 10 ms-nyi, az első formánsok esetében 50 Hz-nél, a második formánsoknál 100 Hz-nél volt kisebb a (statisztikailag szignifikáns) különbség.

A statisztikai adatok részletei ezért értelemszerűen módosítják a fenti kijelentést. Eltekintve a konkrét értékektől, az eltéréseket csupán a funkció szerint különböző [j]-időtartamok magyarázzák nagymértékben (42%-ban). A hangkapcsolatnak nincs hatása az időtartamokra, s bár a 'szóhatár' tényező szignifikáns eredményt igazolt, csak 4,6%-ban magyarázza az itt kapott különbségeket. Mindhárom tényezőt figyelembe véve, az időtartam-különbség szignifikáns [$F(3, 104) = 4,734; p = 0,004; \eta^2 = 0,129$], és közel 13%-ban ad magyarázatot erre az eredményre.

Még mindig eltekintve a konkrét értékektől, az első formánsok nem mutatnak szignifikáns különbséget a funkció függvényében, a 'hangkapcsolat' és a 'szóhatár' tényezők esetében azonban igen; az előbbi 15,4%-ban, az utóbbit 10,2%-ban magyarázzák az adatok. A második formánsoknál a különbség

szignifikáns a funkció és a 'szóhatár' tényező függvényében; az eredményeket 10,1%-ban magyarázzák az adatok az előbbi és 13,5%-ban az utóbbi esetében. A hangkapcsolat, illetve a hangsor jellemzőinek ugyanakkor nem volt hatása az F_2 értékeinek alakulására (annak ellenére, hogy itt az értékekben 200 Hz-nél valamivel nagyobb eltéréseket tapasztaltunk). A korrelációelemzések csupán a hiátustöltő időtartama és a megelőző magánhangzó időtartama között mutattak említésre méltó összefüggést. A V[j]V hangsorok teljes időtartamát tekintve a mássalhangzó aránya mintegy harmada a magánhangzókénak a funkciótól függetlenül. Úgy tűnik tehát, hogy a fonémarealizációk és a hiátustöltők időtartamukban és a második formánsukban különböznek egymástól matematikailag.

Mindezek alapján a konklúzió megfogalmazása mégis egyfajta paradoxon, hiszen a konkrét akusztikai adatok különbségei nagyon kicsik, ennél fogva a vizsgált mássalhangzó minőségi különbsége a funkció függvényében nem meggyőző. A hipotézis tehát egyfelől igazolódott, másfelől azonban azt mondhatjuk, hogy nem. Valószínűtlen, hogy a kapott idő- és formánskülönbségek észlelhetők lennének, ennek igazolása azonban további kutatásokat igényel. Megállapításaink nyilvánvalóan csak a vizsgált anyagra és a mondatokat felolvasókra tekinthetők igaznak. Az adatok – véleményünk szerint – nem támasztották alá egyértelműen, hogy a hiátustöltő [j] mássalhangzó artikulációja különbözne a palatális approximánsétól. A kapott értékek (a szignifikáns különbségek ellenére) nem teszik lehetővé kétféle minősítésüket, vagyis valamennyi [j]-t palatális approximánsnak tekintjük (amelyek időtartama és formánsszerkezete mindkét funkcióban széles tartományban változik). Az időtartam rövidülése önmagában nem belső jellemzője egy adott beszédhang osztályozásának (Maddieson 2008). Korábbi kutatási eredményeink szerint a /j/ fonéma realizációinak formánsszerkezete nagyobb eltéréseket mutatott a kontextus függvényében a jelen adatoknál (Gósy 2014). Ekkor az F_1 esetében az átlagértékek különbsége meghaladta az 50 Hz-et, a második formánsoknál pedig a 100 Hz-et; mégsem állíthatjuk, hogy ezek a mássalhangzók nem ugyanannak a /j/-nek a megvalósulásai lettek volna.

Az összevetések a szakirodalomban korábban közölt adatokkal bizonytalanok, elsősorban a módszertani különbözőségek miatt. Olaszky kísérletében nagyobb különbséget talált a hiátustöltők és a /j/ realizációinak időtartama között, mint a jelen kutatás értékei (2010). A szerző adatai szólista felolvasásán alapultak, számos szempontot vett tekintetbe a kilenc vizsgálati csoport kialakítása során. Valamennyiben különböző magánhangzók szerepeltek az [i] mellett, ugyanazon csoportban előfordultak szó belsejében és szóhatáron is, az összehasonlítás nemegyszer az ugyanazon szóban (eltérő pozíciókban) ejtett hiátustöltő és /j/-realizáció időtartama között történt (Olaszky 2010). Az általunk vizsgált hangsoroknak megfelelő átlagok nála a hiátustöltők esetében 39,1 ms (a négyféle hangkapcsolat átlagából számolva), a fonémarealizációké pedig 57,4 ms (vö. Németh–Olaszky 2010). A VV és a V[j]V hangsorok szá-

ma lényegesen kevesebb volt, mint a mi kutatásunkban (a becsült előfordulás 2 és 7 közötti lehetett). A mérési elvek is különböztek, Olaszynál a hiátustöltő szegmentálása alapvetően egyszemélyes percepciók döntés eredménye volt (2010). Megállapítja továbbá, hogy mind a hiátustöltő, mind a fonémarealizáció időtartama számos tényező következményeként jön létre, s adatai alá is támasztják az időértékek nagymértékű átfedéseit. Mindezek miatt nem adhat egyértelmű választ arra a kérdésre, hogy vajon a hiátustöltő és a fonémarealizáció funkciókülönbsége miként igazolható a mért időtartamokkal.

Menyhárt kutatási anyaga szintén tartalmazta az általunk is vizsgált VV hangkapcsolatokat is (2006), ő ugyancsak mondatokban vizsgálta a hiátustöltők időtartamát és első három formánsát. Kutatási célja azonban más volt, mint a jelen vizsgálatsorozaté. Az ott előforduló (a miénkkel megegyező VV kapcsolatot tartalmazó) mindössze négy szó adatai ismételtelen nem teszik lehetővé, hogy a jelen kutatásban kapott értékekkel összevessük őket. A közölt értékek egyébként jellemzően a saját megfelelő értékeink szórásartományába esnek.

További kutatások adhatnak választ a szógyakoriság esetleges hatására, avagy arra, hogy a tágabb kontextus befolyásolja-e az ejtést (Olaszy 2010), illetve hogy mikor jelennek meg a szokásostól eltérő artikulációs stratégiák (Markó 2012), továbbá az egyéni ejtési sajátosságokra (Cole et al. 2010). Különösen fontos lenne a beszédstílus hatásának ismerete (pl. Aguilar 1999), főként az, hogy vajon a spontán beszédben előforduló hiátustöltők és fonémarealizációk ejtése mikor és miben azonos, hasonló, esetleg eltérő. Annál is inkább, hiszen nem tudhatjuk, hogy a felolvasott szavak, mondatok, szövegek írásképe (ti. *ai* vagy *aji*) milyen mértékben befolyásolja az „írástudó” beszélő kiejtését. Igaz ugyan, hogy ennek az anyanyelvi tudásnak a hatása spontán közlések esetén is befolyásolhatja az adott artikulációs gesztusokat.

Noha több feltételezés ismeretes arról, hogy vajon a beszélők miért ejtenek egy további hangot két szomszédos magánhangzó között (számos nyelvben), a legvalószínűbb magyarázat mégiscsak az lehet, hogy a beszélő törekszik a könnyebb kiejtésre. Az is valószínűsíthető, hogy a beszélő olyan hiátustöltőt alkalmaz, amely tagja az anyanyelve beszédhangállományának. Ez kedvezően hat az artikulációs gesztusok begyakorlottságára, tehát könnyebb kivitelezésére. Elemzéseink nem adtak támpontot arra vonatkozóan, hogy a hiátustöltő jelentősen rövidebb ejtése mivel lett volna magyarázható. Az a tény, hogy jelentős akusztikai eltérést nem találtunk a [j] kétféle funkciójában, arra utal, hogy a kiejtésük nagyon hasonló. A beszélő tehát nem igyekszik a funkcionális különbség megvalósítására azzal, hogy különbözőképpen ejti a [j] mássalhangzókat.

Irodalom

- Aguilar, Lourdes 1999. Hiatus and diphthong: Acoustic cues and speech situation differences. *Speech Communication* 28. 57–74.
- Balassa József 1904. *Magyar fonétika. A hangok és a beszéd fiziológiai elemzése*. Franklin-Társulat, Budapest.
- Beddor, Patrice S. – Harnsberger, James D. – Lindemann, Stephanie 2002. Language-specific patterns of vowel-to-vowel coarticulation: acoustic structures and their perceptual correlates. *Journal of Phonetics* 30. 591–627.
- Bell-Berti, Fredericka – Harris, Katherine S. 1982. Temporal patterns of coarticulation: lip rounding. *Journal of the Acoustical Society of America* 71. 449–454.
- Benguerel, André-Pierre – Cowan, Helen A. 1974. Coarticulation of upper lip protrusion in French. *Phonetica* 30. 41–55.
- Boersma, Paul – Weenink, David 2013. *Praat: doing phonetics by computer [Computer program]. Version 5.3*. <http://www.praat.org>. (Letöltés ideje: 2013. júl. 9.)
- Casali, Roderic F. 1996. *Resolving hiatus*. PhD-dissertation. University of California, Los Angeles.
- Casali, Roderic F. 1998. *Resolving hiatus*. Garland, New York–London.
- Cole, Jennifer – Linebaugh, Gary – Munson, Cheyenne M. – McMurray, Bob 2010. Unmasking the acoustic effects of vowel-to-vowel coarticulation: A statistical modeling approach. *Journal of Phonetics* 38. 167–184.
- Elekfi László 1992. *A magyar hangkapcsolódások fonetikai és fonológiai szabályai*. Linguistica series A, Studia et Dissertationes 10. MTA Nyelvtudományi Intézet, Budapest.
- Farnetani, Edda – Recasens, Daniel 1999. Coarticulation models in recent speech production theories. In Hardcastle, William J. – Hewlett, Nigel (eds.): *Coarticulation: theory, data and techniques*. Cambridge University Press, Cambridge. 31–65.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- Gósy Mária 2014. A kontextus hatása a palatális közelítőhangra. In Havas Ferenc – Horváth Katalin – Kugler Nóra – Vladár Zsuzsa (szerk.): *Nyelvben a világ Tanulmányok Ladányi Mária tiszteletére*. Tinta Kiadó, Budapest. 169–178.
- Haader Lea – Papp Zsuzsanna 1999. *Zsoltároskönyv. Kulcsár-kódex. 1539*. MTA Nyelvtudományi Intézet, Budapest.
- Hardcastle, William 2006. Coarticulation. In Brown, Keith (ed.): *Encyclopedia of Language and Linguistics*. Elsevier, Oxford. 501–505.
- Heid, Sebastian – Hawkins, Sara 2000. An acoustical study of long-domain /r/ and /l/ coarticulation. In *Proceedings of the fifth seminar on speech production: Models and data*. Kloster Seeon, Bavaria, Germany. 77–80.
- Kabak, Barış 2007. Hiatus resolution in Turkish: An underspecification account. *Lingua* 117. 1378–1411.
- Kassai Ilona 1998. *Fonetika*. Nemzeti Tankönyvkiadó, Budapest.
- Kelso, Scott, J. A. – Saltzman, Elliot, L. – Tuller, Betty 1986. The dynamical perspective on speech production: Data, and theory. *Journal of Phonetics* 14. 29–59.
- Maddieson, Jan 2008. Glides and gemination. *Lingua* 118. 1926–1936.
- Manuel, Sharon Y. 1999. Cross-language studies: relating language-particular coarticulation patterns to other language-particular facts. In Hardcastle, William J. –

- Hewlett, Nigel (eds.): *Coarticulation: Theory, data and techniques*. Cambridge University Press, Cambridge, UK. 179–198.
- Markó, Alexandra 2012. Boundary marking in Hungarian V(♯)V clusters with special regard to the role of irregular phonation. *The Phonetician* 105–106. 7–26.
- Menyhárt Krisztina 2006. Koartikulációs folyamatok két magánhangzó kapcsolatában. *Beszéd kutatás* 2006. 44–56.
- Németh Géza – Olasz Gábor (szerk.) 2010. *A magyar beszéd. Beszédkutatás, beszédtechnológia, beszédinformációs rendszerek*. Akadémiai Kiadó, Budapest.
- Odden, David 1996. *The phonology and morphology of Kimatumbi*. Oxford University Press, Oxford.
- Odden, David 2006. Kimatumbi: Phonology. In Brown, Keith (ed.): *Encyclopedia of Language and Linguistics*. Elsevier, Oxford. 198–206.
- Olasz Gábor 2006. *Hangidőtartamok és időszervezeti elemek a magyar beszédben*. Nyelvtudományi Értekezések 155. Akadémiai Kiadó, Budapest.
- Olasz Gábor 2010. Az [i] + V és V + [i] hangkapcsolódások akusztikai elemzése a hiátustöltés magyarázatához. *Beszéd kutatás* 2010. 76–81.
- Öhman, Sven E. G. 1966. Coarticulation in VCV utterances: Spectrographic measurements. *Journal of the Acoustical Society of America* 39. 151–168.
- Padgett, Jaye 2008. Glides, vowels, and features. *Lingua* 118. 1937–1955.
- Papp István 1971. *Leíró magyar hangtan*. Tankönyvkiadó, Budapest.
- Quintilianus, Marcus Fabius 2008. *Szónoklattan*. Kalligram Kiadó, Pozsony.
- Siptár Péter 2001. Egy jottányi fonológia. *Beszéd kutatás* 2001. 1–17.
- Siptár Péter 2002. Hiátus. In Hunyadi László (szerk.): *Kísérleti fonetika, laboratóriumi fonológia*. Debreceni Egyetem Kossuth Egyetemi Kiadója, Debrecen. 85–98.
- Siptár, Péter – Törkenczy, Miklós 2000. *The phonology of Hungarian*. Oxford University Press, Oxford.
- Szende Tamás 1992. *Alapalak és lazítási folyamatok*. MTA Nyelvtudományi Intézet, Budapest.
- Szigetvári Péter 1998. Kormányzás a fonológiában. *Általános Nyelvészeti Tanulmányok* 19. 165–213.
- Uffmann, Christian 2007. Intrusive [r] and optimal epenthetic consonants. *Language Sciences* 29. 451–476.
- Vértes O. András 1982. A beszédhangok akusztikai vetülete. In Bolla Kálmán (szerk.): *Fejezetek a magyar leíró hangtanból*. Akadémiai Kiadó, Budapest. 71–114.

MAGÁNHANGZÓ-IDŐTARTAM A KÖVETŐ MÁSSALHANGZÓ ZÖNGÉSSÉGÉNEK FÜGGVÉNYÉBEN

Grácz Tekla Etelka

Bevezetés

Beszédünk során az egyes fonémák képzési jegyei hatnak egymásra, az így létrejövő beszédhangok ezért fonetikai környezetükkel kölcsönhatásban más-más módon, kisebb-nagyobb eltérésekkel realizálódnak (Gósy 2004). Viszonylag univerzális sajátossággként szokás említeni, hogy a magánhangzók időtartama a zöngés mássalhangzók előtt hosszabb, mint a zöngétlenek előtt. Az egyes nyelvek között ugyanakkor kisebb-nagyobb eltéréssel érvényes ez a tendencia, angolban találták a legerősebbnek (Maddieson 1997). A magyar nyelven végzett eddigi vizsgálatok azt mutatták, hogy a hatás megjelenése függ a magánhangzótól magától és a követő mássalhangzótól is (vö. pl. Olasz 2000; Kovács 2002; Bárkányi-Kiss 2009; Grácz 2012). A kutatások ugyanakkor többségében kevés (gyakran egy-egy) beszélő elemzésével készültek, így általános megállapítás levonásához további vizsgálatok szükségesek.

Feltételezett magyarázatok

A jelenség hátterében több magyarázatot is feltételeznek (összefoglalóan Maddieson 1997). Az egyik az artikuláció és az aerodinamika összefüggéseire épít, miszerint a zöngétlen obstruensek képzése alatt nagyobb az intraorális nyomás (a hangszalagok felőli szabad levegőáramlás következtében), ezért feszesebb, intenzívebb akadályképzést feltételeznek ezeknél. Az elmélet szerint ehhez gyorsabb mozgás szükséges, ami miatt a célt előbb éri el a beszédképző szervek, azaz a magánhangzóból a mássalhangzóba való átállás előbb történik meg a zöngétlenek esetében (Chen 1970). Ennek az elképzelésnek több kutatás eredménye is ellentmond, legalábbis egyedüli érvényességét cáfolja. Egyrészt a homorgán szonor mássalhangzók is hosszabbak a zöngés mássalhangzó előtt (pl. *nt*, *nd*; Raphael et al. 1975). Másrészt felpattanók esetében azt találták, hogy nem csak előbb kezdődik a zárképzés, hanem a zár felnyílása is később kezdődik (Dunn 1993), emiatt nagyobb az opozícióban álló mássalhangzók időtartamának eltérése, mint a megelőző magánhangzóké.

Maddieson (1997) másik magyarázatként felveti, hogy a hasonló hangsorok időtartamának kiegyenlítése is magyarázhatja a rövidebb mássalhangzó előtti hosszabb hangzóképzést.

A percepció hatása kétféle elmélet alapján is felmerül. Javkin (1976) feltételezi, hogy a zöngének a mássalhangzó ideje alatti további fennmaradása a percepcióban a magánhangzó hosszabb észleletét kelti, aminek következtében a nyelvelsajátítás során ezt követendő időtartam-különbségként sajátítjuk el. Más elmélet szerint a hosszabb magánhangzó-időtartam hatására a mássalhangzót rövidebbnek észleljük, azaz a zöngés párokat megelőző hangzók eltérése a kontraszt kiemelését célozza.

Feltehetően az egyes magyarázatok közül nem egy, hanem azok együtt, kölcsönhatásban okozzák a kérdéses időtartam-mintázatot.

Korábbi eredmények magyar nyelvre

Magdics (1966) és Kassai (1979) eredményei szerint a magánhangzók időtartama hosszabb zöngés mássalhangzók, mint zöngétlen párjuk előtt. Olasz (2000) néhány kivételt talált. Az *e* felpattanók előtt nem követte ezt a tendenciát, és az *ö* is kivételként jelent meg. Az *a* és az *u* esetében találta a legnagyobb hatást. A réshangok előtt nagyobb eltérést talált, mint a felpattanók esetében. Ez utóbbit Kovács (2002) is igazolta, de kiemelte, hogy a megelőző hangzó is jelentős hatást gyakorol. Kohári (2010) ugyancsak talált a magánhangzó és a mássalhangzó jegyei alapján is eltérést a tendencia érvényesülésében: az *ö* és *ü* esetében explozívák előtt, az *o* pedig az alveoláris réshangpárral alkotott kapcsolatban nem követte azt.

Bárkányi és Kiss (2009) vizsgálatában két magánhangzót és két réshangpárt elemzett. A mássalhangzó szó végi helyzetében az *é sz* és *z* előtt jól elkülöníthető volt az egyes beszélők ejtésében, az *á* időtartamában *s* és *zs* előtt nagyobbak voltak az időtartamszórásban az átfedések az egyes beszélők ejtésében. A mássalhangzó pár időtartama nagyobb eltérést mutatott mindkét esetben, de jellemzően azon beszélőknél, akiknél kisebb volt az eltérés a V-időtartamban, a C-ében is. A szó belseji helyzetben az *é* kisebb szóródást és magasabb átlagot mutatott a *z* előtt, de teljes tartománya a *sz* előtti vokális időtartamának tartományával átfedett, egy beszélő ejtésében pedig ellentétes tendenciát adatoltak. Az *á* három beszélőnél egyértelműen mutatta a várt tendenciát, a másik három esetében változó módon eltért ettől. Ebben a pozícióban a konzonáns időtartama is nagyobb átfedéseket mutatott, de a tendencia egyértelmű volt a legtöbb beszélőnél.

Grácsi (2012) vizsgálatában az *a* magánhangzót elemezte a fonetikai helyzet függvényében. A magánhangzók időtartama az explozívák előtt csak abszolút szóvégen a bilabiális és a veláris pár előtt mutatott eltérést, a réshangok esetében a *Calal* hangsorban mindhárom pár előtt, szó belseji (*laCal*) helyzetben csak a posztalveoláris előtt, szó végi helyzetekben pedig a szibilánsok előtt követte a tendenciát. A magánhangzók közötti eltérést szó belsejében három vokális időtartamán elemezte (*a, i, u; IVCVI* hangsorban). A réshangok előtt az *a* csak a posztalveolárisok előtt, a két másik magánhangzó mindhárom pár előtt követte a tendenciát, az explozívák előtt nem volt kimutatható szignifikáns eltérés.

Összegezve azt mondhatjuk, hogy a hatást nem minden tanulmány mutatta ki egyértelműen, amelyek igen, azokban is jellemzően néhány magánhangzó és/vagy a mássalhangzó eltért attól, illetve a fonetikai helyzet alapján is változtak az eredmények. Ennek hátterében több okot is feltételezhetünk. Az egyik, hogy nem minden vizsgálat készült minimálpárok alkalmazásával, ezért a későbbiekben a hangkörnyezet szerepe is elemzendő. Másrészt szó végi pozícióban, főleg abszolút szó végiben jelentősebb a neutralizálódás a fonetikai zöngésségben, ezért feltételezhető, hogy ha a magánhangzó-időtartam alakulásának (egyik) meghatározója a percepció segítése, akkor a mássalhangzó fonetikai zöngésségének megtartására kedvezőtlenebb helyzetekben a vokális kiemeltebb szerepet játszik, így a kérdéses hatás is erőteljesebb (Bárkányi–Kiss 2009). Ezt alátámaszthatja az is, hogy az explozívák esetében kisebb arányú és ritkább a zöngétlenedés, de hangsor végén jelentős, és hogy általában szóvégen a C-időtartam is erőteljesebben eltér. Ugyanakkor az időtartamok eltérése az explozívák esetében is nagyobb szóvégen. Az egyes beszédhangok eltérően érintettek a hangsor végi nyúlásban (pl. Grácz 2012).

Az sem hagyható figyelmen kívül, hogy a hangsorbeli helyzettel a szótag szerkezete, esetleg a magánhangzó hangsorbeli egyéb jellemzői is (első vs. többedik szótag) is megváltozhatnak. A hangsúlyos és hangsúlytalan szótagok összevetése ellentétes eredményeket hozott korábban (pl. Magdics 1966; Kassai 1982), spontán beszédben igazolták a magánhangzó hosszabb voltát első szótagban (Gósy–Beke 2010). Az egyes nyelvekben a hangsor végi nyújtás nem azonos elemszámig hat előre, magyarra még pontos adatok nincsenek. Kérdés, hogy hány szegmensig hat, és ez befolyásolhatja-e és hogyan az egyes elemzett pozíciókat (Fletcher 2010). A vizsgált magánhangzó az elemzésekben állhatott (C)V szótagban a hangsorbeli helyzetekben, míg szóvégen szükségszerűen CVC-ben állt. Felvethetjük kérdésként a mássalhangzó hatókörét a szótag függvényében.

Magyarázként felvethető, hogy a magánhangzó valamely jellemzőjével függ össze a tendenciának ellentmondó eredmények előfordulása. Kérdés például, hogy a hosszúsági oppozíció megőrzése végett korlátozott lehet-e az időtartamok változékonysága (vö. Keating 1984 felvetését a lengyelre). Az egyes vizsgálatokban középső vagy alsóbb nyelvéllátásfokú magánhangzók „sértették meg” a tendenciát, ugyanakkor az *a* és *e* esetében kevésbé releváns az időtartam szerepe a megkülönböztetésben. A kvantitásbeli párok spontán beszédben minden esetben eltértek, a hangkörnyezettől függetlenül – Gósy–Beke 2010; más vizsgálatokban megkérdőjeleződött az időtartam szerepe (pl. Mády–Reichel 2007). Mindkét eredmény ellentmond a fenti feltételezésnek, az azonban figyelembe veendő, hogy milyen mértékűek a hosszúsági párok időtartamának átfedései és a mássalhangzó hatása. Ugyanakkor az egyes beszédhangok észlelését hangkörnyezetben sajátítjuk el, így a hangkörnyezetre jellemző tendenciák szerint kódolódik az akusztikai lenyomat (Ohala 1983). A vokálisok artikulációs konfigurációjuk következtében jellemző időtartam-

beli sorrendbe állíthatóak (van Santen 1992), a legrövidebbek kevésbé nyomhatóak össze (Nooteboom 1997). Ismét ellentmond ennek a lehetséges magyarázatnak, hogy az *i* és *u* a legtöbb tanulmányban követte az elvárt tendenciát, amennyiben a kérdéses hatás háttérében nem nyúlást feltételezünk. Ugyanakkor a követő mássalhangzók képzési helye alapján is található eltérés, nem csak a magánhangzók képzési jegyei mentén, tehát a fonetikai távolság hatása feltételezhető. Az artikulációs pontok közötti váltás időzítése és aerodinamikai tényezők is szerepet játszhatnak az eredményekben. A magánhangzó konfigurációjából a mássalhangzóéba történő átállás folyamatos, a szupraglottális és a glottális mozzanatok időzítése eltér a zöngésségi párok között. A zöngés képzés során a cél a zöngét és a zárat/zörejt egyszerre képezni, az explozívák ejtésében legfeljebb a szájüregi nyomás miatt áll le a zöngé, a beszélő kompenzációs stratégiákat alkalmaz(hat) a fenntartás érdekében (Ladefoged–Maddieson 1996). Ezzel szemben a réshangok esetében a zöngé képzése során alacsonyabb a szájüregbe érkező légáram, így a frikatíva nem lehet elég intenzív (Stevens 1998), ezért a zöngé „zavaró”, nagyobb a zöngétlenedés aránya és valószínűsége. Zöngétlen képzés során az explozívák ejtésekor nem a zöngé mielőbbi leállítása az artikulációs cél, hanem a fenn nem tartása (= passive devoicing; Ladefoged–Maddieson 1996) a frikatívákkal szemben, ahol a zöngé mielőbbi leállítása a cél (= active devoicing; Stevens 1998). A magánhangzó létrehozásához alacsony szájüregi nyomás és a zöngé jelenléte szükséges. Az egyes vokálisok konfigurációjából az egyes konzonánsokhoz való átmenethez eltérő „hosszúságú” utat tesznek meg a képzőszervek, ezért a különböző hangzók létrehozásához szükséges aerodinamikai feltételek eltérő időzítéssel épülhetnek le és állhatnak fel. Mindezek mellett az a lehetőség is felmerül, hogy a beszéd- és artikulációs tempót nem kontrollálják az egyes kísérletek, és így az egyes hangsorok közötti eltérések esetleg lehetnek olyan mértékűek, hogy elfedik a magánhangzók egyes helyzetek közötti eltérését. Bárkányi és Kiss (2009) vizsgálatában is jellemző, hogy a V : C időtartamarányok jobban elkülönülnek a zöngésségi párok előtt, mint a magánhangzó-időtartamok. Meg kell azonban jegyezni, hogy a mássalhangzók időtartama az arányítás nélkül is konzekvensebb eredményeket mutatott – ahogyan azt korábban is említettük.

A jelen kísérletet előtanulmánynak szánjuk, első lépésnek a feltételezhető magyarázatok elemzésében. Két mássalhangzó párt vizsgáltunk, az alveoláris explozívákat és réshangokat. Főbb kérdéseink, hogy (i) mely magánhangzók esetében érvényesül a tendencia, hogy a zöngésségi oppozíciót alkotó mássalhangzó párok zöngés tagja előtt hosszabb időtartamban valósulnak meg, mint a zöngétlen előtt? (ii) Milyen szerepe van a képzési módnak? (iii) Milyen hatással van a hangsorbeli pozíció? (iv) Mennyiben beszélőfüggő a tendencia megjelenése?

Kísérleti személyek, anyag és módszer

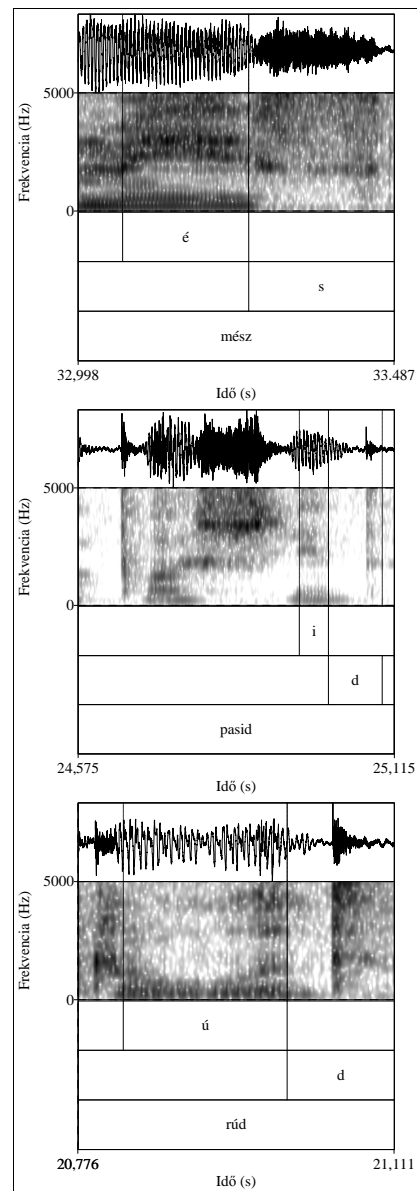
A kérdések megválaszolásához 6 adatközlő [3 nő (n1, n2, n3) és 3 férfi (f1, f2 és f3)] ejtésében rögzítettünk olvasott mondatokat. A beszélők a budapesti köznyelvet beszélik, sem hallási, sem beszédhibájuk nem ismert. Életkoruk 24 és 30 év között volt.

A felolvasandó anyagban a 14 magyar köznyelvi magánhangzó szerepelt a /d t z s/ előtti megelőző helyzetben, utolsó és utolsó előtti szótagban. Törekedtünk arra, hogy (i) a célszavak két szótagúak legyenek, (ii) a vizsgált magánhangzó előtti fonéma azonos legyen az azonos képzésmódú mássalhangzók előtt (lehetőleg minimálpárt alkossanak a célszavak), illetve (iii) hogy a zöngésségi párok tagjai esetében a hordozó mondat prozódiaileg hasonló szerkezetű és azonos szótagszámú legyen. Az adott helyzetben ritkább hangkapcsolatok esetében néhányszor szükséges volt az első két ponttól eltérő szópárokat alkalmazni (pl. *íszom-izom*, *csúszás-húzás*), esetleg logatomokat használni (pl. *bűz-BŰSz*, *sóz-SÓSz*). Ez utóbbiak esetében úgy instruáltuk az adatközlőket, hogy lesz néhány mondatban egy-egy társaságról is szó, pl. a Svájci Órások Szövetségéről. A mondatokat randomizált sorrendben háromszor olvasták fel az adatközlők. Beszélőnként 336, összesen pedig 2016 mondatot rögzítettünk. A felolvasásokat csendesített szobában AT-4040 mikrofonnal számítógépre rögzítettük. Az adatközlők annyiszor állhattak meg pihenni, ahányszor szükségét érezték.

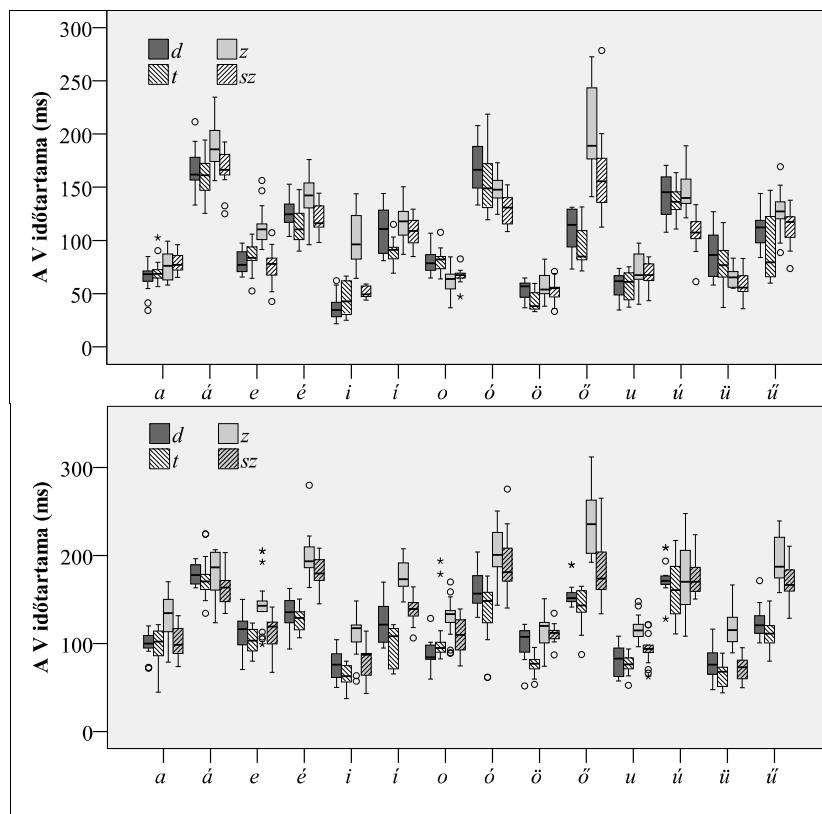
A magánhangzók és a követő mássalhangzók időtartamát, valamint azok arányát elemeztük (1. ábra). A magánhangzó második formánsának megindulását és lecsengését vettük figyelembe. A méréseket a Praat 5.3 szoftverrel (Boersma–Weenink 2013) végeztük.

Eredmények

Az egyes magánhangzók időtartama a kétszer négy elemzett környezetben változatosan alakult (2. ábra). Az elvárt mintázatot, miszerint a mássalhangzó-párok zöngés tagja előtt hosszabb, a zöngétlen előtt rövidebb ugyanazon magánhangzó időtartama, viszonylag kevés esetben követték az adatok. Az átlagos időtartamok között a független mintás *t*-próba alapján az explozívák előtt utolsó szótagban hat (*í, ó, ö, ő, ü, ű*), az utolsó előttiben öt (*é, í, ő, ő, ű*) magánhangzó követte a várt tendenciát. A frikatívák előtt az utolsó szótagban tizenegy (*a, á, e, é, i, í, o, ő, u, ü, ű*), az utolsó előttiben nyolc (*á, e, é, i, ő, ö, ú, ű*) esetben találtunk ilyen eltérést. A felpattanókat megelőző pozíciókban az utolsó előtti szótagban az *o* nem szignifikáns (19,1%), az utolsóban az *i* időtartama szignifikánsan (24,5%) a zöngétlen konsonáns előtt volt hosszabb.



1. ábra
Példa a címkézésre



2. ábra

A magánhangzók időtartama az utolsó előtti szótagban (fent) és az utolsó szótagban (lent) a négy mássalhangzó előtt

Az összes beszélő ejtésének közös vizsgálatában tehát az állapítható meg, hogy a két képzési mód közül a réshangok esetében nagyobb mértékben jellemző a várt tendencia megjelenése, illetve hogy az utolsó előtti szótagban több vokális követi azt. A képzési módok között ez az eltérés várható a korábbi, magyar mássalhangzókra vonatkozó irodalom eredményei alapján, mivel a réshangok gyakrabban és nagyobb arányban zöngétlenednek, és ezzel párhuzamosan a mássalhangzók időtartama is nagyobb mértékben követi a várt tendenciát, miszerint a zöngés konzonáns rövidebb zöngétlen párjánál. A szótagok helyzete alapján kapott eredmények ugyancsak megfelelnek a korábbi adatoknak a mássalhangzók zöngésségi és időtartam-eredményeiről.

A jelen vizsgálatban a zöngés mássalhangzók átlagos időtartama minden magánhangzót követő helyzetben szó belsejében 19,4–36,4%-kal, szóvégen 24,9–48,1%-kal rövidebb volt, mint zöngétlen párjuké. Az eltérés a független mintás *t*-próba szerint minden esetben szignifikáns volt. A magánhangzó és a mássalhangzó időtartamának arányát vizsgálva minden zöngésségi pár tagjai között 17–66% eltérés tapasztalható. A független mintás *t*-próba szerint az utolsó szótagban az *-od/-ot*, az utolsó előttiben az *-id/-it* esetében nem, a többi hangkapcsolatban szignifikáns az eltérés. A veláris magánhangzó esetében a két mondat csak egy fonémában tért el (*A kutya érzi a szagot/szagot.*), a palatális esetében azonban a mondatok jobban eltértek (*Ez a kóla nem is hideg./Elvesztett már minden hitet.*). Mivel nagyobb eltérés több mondatpárnál is előfordult, ezt csak egyik okként feltételezhetjük a jelenség hátterében.

Az egyes adatközlők ejtését külön-külön elemezve nagy beszélők közötti eltéréseket kaptunk (3. ábra). Az adatközlők száma természetesen nem teszi lehetővé, hogy ezen variabilitás alapján végleges következtetést vonhassunk le, azonban a nagy különbségek létezésére felhívja a figyelmet, illetve a jelen adatok értelmezéséhez elengedhetetlen az elemzésük. A vizsgálat során azokat a magánhangzókat tekintettük a várt tendenciát követőnek, amelyek esetében a beszélő ejtéseinek a zöngétlen mássalhangzó előtti átlagos magánhangzó-időtartam legalább a 10%-kal alacsonyabb volt, mint a zöngés előtt, ennek ellentmondónak pedig azokat, amelyeknél ez az érték legalább 10% magasabb volt a zöngés konzonáns előtti értéktől, tehát az eltérés –10% vagy annál kisebb volt. Az utolsó szótagban csak egy beszélőnél volt több mint a magánhangzók fele (8) hosszabb a zöngés felpattanó előtt, mint a zöngétlen előtt, míg a réshangok esetében öt beszélőnél. Legalább egy olyan magánhangzót négy beszélőnél találtunk a felpattanók esetében, ahol a várt eltérés fordítottja jelentkezett, a réshangoknál pedig egy beszélőnél (f2), akinek az ejtésében az explozívák esetében is gyakoribb volt ez a jelenség. Az utolsó előtti szótagban a réshangok előtti magánhangzókra ismét azt az eredményt kaptuk, hogy a hatból öt beszélő ejtésében követte legalább nyolc a várt tendenciát, az egy beszélő (n3), akinél nem haladta meg az 50%-ot, eltért az utolsó szótagnál hasonló eredményt adó beszélőtől (f2). Nem volt olyan beszélő, akinek a produkciójában legalább nyolc magánhangzó követte a tendenciát az explozívák előtt. Ebben a pozícióban minden beszélő ejtésében volt olyan magánhangzó, amelyik jellemzően hosszabb volt a zöngétlen explozíva előtt, mint a zöngés előtt; a réshangok esetében pedig ez csak a három férfira volt jellemző.

Az adatok tehát ismét igazolják a képzési módok között és a fonetikai helyzetek közötti eltérést. A beszélők közötti variabilitás ezen tényezőkkel együttesen jelenik meg.

A mássalhangzók időtartamát a fenti módon elemezve kevesebb esetben találtunk saját zöngétlen párjánál hosszabb zöngés mássalhangzót. Ezen ese-

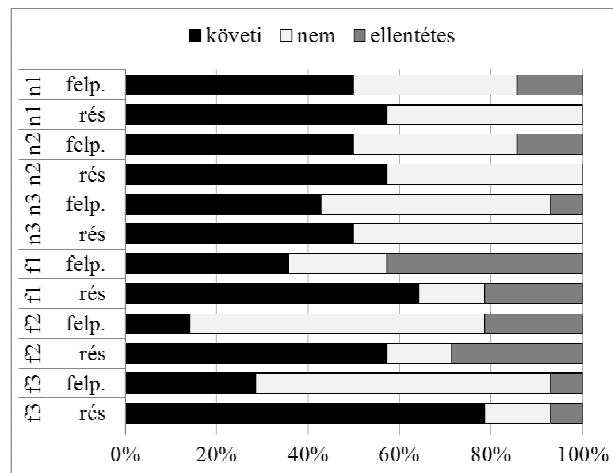
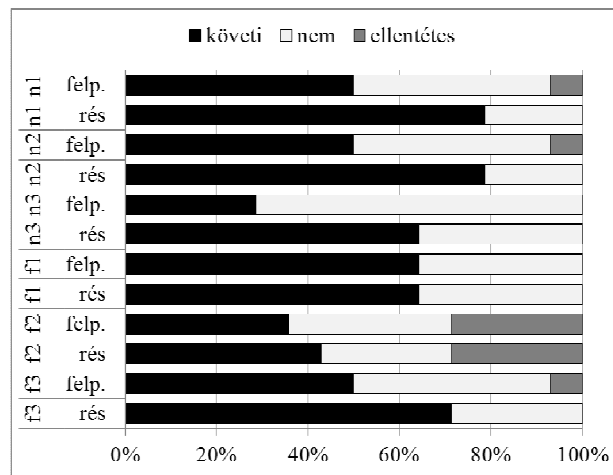
tekben is megállapítható volt a beszélők közötti eltérés. Egy nőnél (n3) és mindhárom férfinál megjelent a tendenciát nem követő mássalhangzó. Míg a nő esetében az *-iz#/-isz#* hangsorok ejtése, addig a férfiaknál gyakrabban a *-Vd#/-Vt#* hangsoroké mutatott hosszabb zöngés mássalhangzót. Ebben az esetben is f2 produkciója tér el leginkább az elvárttól. Míg f1 két, f3 három szó végi fonémapár megvalósításában, addig f2 hat szó végi és két szó belseji konzonáns esetében ejtette rövidebben a tag zöngétlen párját. Nem találtunk olyan összefüggést, hogy ha egy adott magánhangzó nem követte a beszélő ejtésében a várt tendenciát, akkor az azt követő mássalhangzók sem, avagy fordítva. Tehát nem az egyes hangsorokra, hanem a beszélő ejtésében az adott hangra tűnik jellemzőnek a várt szabályosságnak való meg nem felelés.

A magánhangzó-mássalhangzó időtartamarány esetében két beszélő, n3 és f1 esetében nem találtunk olyan hangsort, ahol nem a zöngés konzonánst tartalmazó hangsor V : C időtartam-aránya lett volna nagyobb. Három beszélő, n1, n2 és f3 esetében egy-egy ilyen hangsortpár volt: *-ed/-et* (n1, n2) és *-űz/-űsz* (f3). Az egyik férfi, f2 produkciójában ismét több hangsortpár volt található, amely nem követte a várt eltérést: hét szó végi és két szó belseji explozívapár, valamint egy szó végi frikatívapár. Csak ez utóbbi beszélő esetében találtunk összefüggést a hangsortpárok egyes paraméterei között: azon szavak ejtésében, amelyeknél a V : C arány nem haladta meg a 10%-ot a zöngés-zöngétlen párok előtt, a magánhangzó időtartama sem követte a várt tendenciát.

A mássalhangzók időtartamára várható tendencia követése is eltérően alakult az egyes beszélők ejtésében. A C-időtartamoknak a nem várt módon alakulása a magánhangzókéval ellentétben a szó belseji helyzetben ritka volt, f2 ejtésében az *é* és *i* utáni felpattanók tértek el a várt eredményektől; szó végén pedig gyakrabban, mindhárom férfi és n3 produkciójában is jelent meg ilyen hangsortpár. Két beszélőnél, n3 és f3 esetében a réshangokat (is) érintette (egy és két hangsortpár esetében), míg f1 és f2 ejtésében csak a felpattanókat (két és hat esetben, amely utóbbiból egy szekvencia a várttal ellentétes eredményt adott). Ugyancsak egy hangsortpárt érintett f3 esetében a felpattanók esetében is.

A V : C időtartamarányok az utolsó szótagi magánhangzó esetében csak egyszer, f3 ejtésében egy hangsortpárnál (*-űz#/-űsz#*) mondtak ellent a várt tendenciának, és hét felpattanós hangsortpár esetében (egy annak ellentétes eredménytel) és egy réshanggal alkotott kapcsolatpárban f2 ejtésében. Hangsor belseji helyzetben az *-ed/-et* két beszélőnél (n1, n2) nem követte a várt tendenciát, illetve további két explozívás hangkapcsolat tért el attól f2 ejtésében.

Elmondható tehát, hogy ugyan az egyes mérések esetében részben eltérően alakulnak az adatközlők közötti különbségek, ugyanakkor akire az egyik esetben nagymértékben vagy egyáltalán nem jellemző a várttól eltérő és/vagy annak ellentmondó ejtés, azok produkciója a másik mérés alapján is hasonlóan „deviáns”/elvárt jellegű. Azaz a beszélők közötti eltérés jelentős hatással van az eredményekre és a levonható következtetésekre.

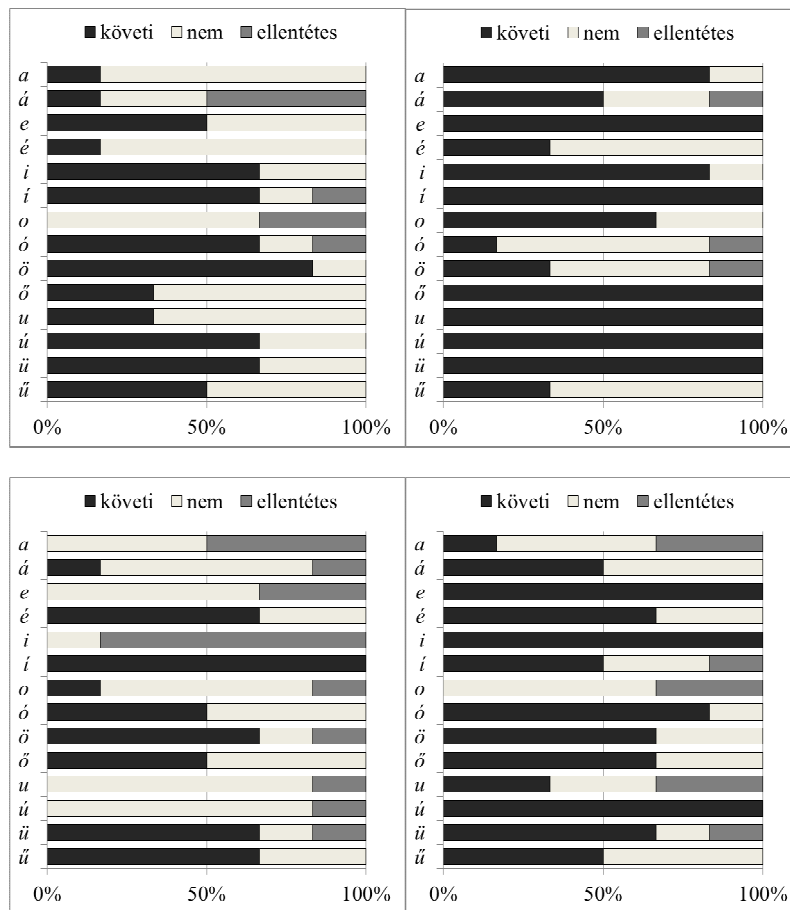


3. ábra

A tendenciát követő és annak ellentmondó magánhangzók eloszlása beszélőnként.

(Fent: utolsó szótag, lent: utolsó előtti szótag. Felp. = felparranók.)

Elemeztük, hogy az adott magánhangzó(k)ra jellemzőnek tekinthető-e, hogy követik, nem követik, esetleg ellentmondanak a várt tendenciának (4. ábra). Ebben az esetben is az átlagok 10%-nál nagyobb eltérését tekintettük úgy, hogy az adott hangsor követte, a –10%-nál kisebbet pedig, hogy ellentmondott az elvárt eredményeknek.



4. ábra

A tendenciát követő és annak ellentmondó magánhangzók beszélőnként.
 (Fent: utolsó szótag, lent: utolsó előtti szótag; balra: felpattanó zárhangok,
 jobbra: réshangok.)

Az utolsó szótagban az explozívák előtt hat magánhangzó (*i, í, ó, ő, ú, ü*) esetében volt jellemző a hosszabb időtartam a zöngés mássalhangzó előtt, a frikatívák előtt pedig kilenc (*a, e, i, í, o, ő, u, ú, ü*) esetében. Négy és három vokális esetben fordult elő, hogy volt olyan beszélő, akinél a zöngétlen előtti ejtés jelentősen hosszabb volt a zöngés előttinél (felpattanók: *á, i, o, ó*, réshangok: *á, ó, ő*). Az utolsó előtti szótagban eggyel kevesebb magánhangzó adta a várt eredményeket legalább négy beszélőnél. Mindkét esetben részben

eltért, hogy mely magánhangzók követték a tendenciát az egyik, melyek a másik pozícióban. Ezek a felpattanók esetében az *é, í, ó, ö, ú, ü*; a réshangok esetében az *e, i, ó, ö, ő, ú, ü* vokálisok voltak. Az is eltérő volt az előző helyzettől, hogy mely magánhangzók időtartamára kaptunk a várttal ellentétes eloszlási eredményeket legalább egy adatközlőnél. A felpattanók esetében kilenc magánhangzónál (*a, á, e, é, i, o, ö, u, ú, ü*), a réshangok esetében csak ötnél (*a, í, o, u, ü*) mértünk ilyen realizációkat.

Mind a mássalhangzók időtartamára, mind a $V : C$ időtartamarányra kapott eredmények esetében minden magánhangzó legfeljebb két adatközlő produkciójában nem követi a várt tendenciát, azaz inkább egy-egy beszélő egy-egy hangsorára jellemző a „deviáns” ejtés, nem pedig a hangsorokban vizsgált magánhangzókkal függ össze. Mindkét esetben több ilyen adatot kaptunk a hangsor végi, mint belseji helyzetben (C időtartama: 12 és 2; $V : C$ arány: 9 és 4). A réshangok esetében pedig csak az előbbiben fordult elő (C időtartama: 3, illetve $V : C$ arány: 2), míg a felpattanókra mindkét helyzetben gyakoribb volt.

Ezek alapján a magánhangzók esetében a hangsorbeli helyzet függvényében meghatározó volt magának a vokálisnak a minősége, míg a mássalhangzók esetében nem, ott a konzonáns saját képzési módja és a hangsorbeli helyzet alapján oszlottak meg az eredmények, a két hangzó időtartamának arányára pedig ugyancsak az utóbbi két tényező gyakorolt hatást.

Az ismételt mérés statisztikai próba alapján (1. táblázat) elmondható, hogy a magánhangzó időtartamát utolsó szótagi pozícióban a követő mássalhangzó zöngéssége és a beszélő ejtési sajátosságai együttesen határozzák meg, míg az utolsó előtti szótagban a vizsgált hangkörnyezetek közül a réshangok esetében igazolódott a zöngésség hatása. A mássalhangzók esetében a korábbi eredmények alapján várhatóan a zöngésség önálló és annak az adatközlő beszédsajátosságaival együttes hatása mutatkozott. Egy esetben, a szó belseji felpattanók ejtésében a magánhangzó-minőséggel való interakció is jelentős volt. A két hangzó időtartamának arányát ugyancsak a zöngésség és annak a beszélők közötti variabilitással közösen gyakorolt hatása befolyásolta mindkét elemzett helyzetben mindkét vizsgált képzési mód esetében. A zöngésség és a magánhangzó-minőség interakciója csak egy elemzett alcsoportban, az utolsó szótagi magánhangzók felpattanókkal alkotott kapcsolatában bizonyult statisztikailag relevánsnak.

1. táblázat: Az ismételt mérés varianciaanalízis eredményei. (Z = zöngésség, V = magánhangzó-minőség; B = beszélő. Dölt = szignifikáns.)

				F	p	r ²
A magánhangzó időtartama	d/t	Utolsó szótagi magánhangzó	Z	1,863	0,176	0,024
			Z * V	1,574	0,213	0,020
			Z * B	2,920	0,018	0,159
		Utolsó előtti szótagi magánhangzó	Z	0,106	0,746	0,001
			Z * V	8,707	0,004	0,102
			Z * B	1,905	0,103	0,110
	z/sz	Utolsó szótagi magánhangzó	Z	28,316	< 0,001	0,269
			Z * V	0,001	0,976	0,000
			Z * B	3,853	0,004	0,200
		Utolsó előtti szótagi magánhangzó	Z	14,823	< 0,001	0,161
			Z * V	0,126	0,724	0,002
			Z * B	1,439	0,220	0,085
A mássalhangzó időtartama	d/t	Utolsó szótagi magánhangzó	Z	132,601	< 0,001	0,633
			Z * V	0,793	0,376	0,010
			Z * B	18,161	< 0,001	0,541
		Utolsó előtti szótagi magánhangzó	Z	141,325	< 0,001	0,647
			Z * V	4,973	0,029	0,061
			Z * B	3,364	0,008	0,179
	z/sz	Utolsó szótagi magánhangzó	Z	146,499	< 0,000	0,655
			Z * V	0,268	0,606	0,003
			Z * B	4,322	0,002	0,219
		Utolsó előtti szótagi magánhangzó	Z	426,348	< 0,001	0,847
			Z * V	0,091	0,763	0,001
			Z * B	7,419	< 0,001	0,325
A V : C időtartamarány	d/t	Utolsó szótagi magánhangzó	Z	72,629	< 0,001	0,485
			Z * V	0,215	0,644	0,003
			Z * B	11,916	< 0,001	0,436
		Utolsó előtti szótagi magánhangzó	Z	27,319	< 0,001	0,262
			Z * V	3,331	0,072	0,041
			Z * B	1,456	0,214	0,086
	z/sz	Utolsó szótagi magánhangzó	Z	162,071	< 0,001	0,678
			Z * V	1,288	0,260	0,016
			Z * B	4,396	0,001	0,222
		Utolsó előtti szótagi magánhangzó	Z	79,753	< 0,001	0,509
			Z * V	1,636	0,205	0,021
			Z * B	0,748	0,590	0,046

Következtetések

A jelen tanulmányban a magánhangzók időtartamának egy befolyásoló tényezőjét elemeztük, a követő mássalhangzó zöngésségét az alveoláris felpattanó zár- és réshangok környezetében. Ennek a hatásnak a meglétét a magyar

nyelvben néhányan elvetik, mások eredményei alapján egy-egy esetben nem érvényesül. Ez utóbbiak azonban kísérletes eredmények, és vagy csak egy-két beszélőt elemeztek, vagy csak egy-két hangkapcsolatot. Ebben a vizsgálatban az összes magyar magánhangzót elemeztük, két mássalhangzópár előtt két pozícióban, hat beszélő ejtésében.

Az eredmények azt mutatják, hogy a magánhangzók időtartamára a követő mássalhangzó hatással van, ez a hatás azonban erőteljesebb a mondat utolsó szótagjában, mint az utolsó előttiben; erősebb a réshangok, mint a felpattanó zárhangok előtt; nagymértékben függ a beszélő ejtésének sajátosságaitól; és feltehetően a magánhangzó és a mássalhangzó jegyeitől is függ.

A pozíció és a konzonáns képzésmódjának hatása a magánhangzó időtartamára várhatóan összefüggésben van azzal, hogy ezek mentén a mássalhangzó zöngétlenedése is eltér, a szóvégen és a réshangok esetében gyakoribb és magasabb arányú. Feltehetően az artikulációs mozzanatok változása (szakasz végi nyújtás) és a konzonáns zöngésségi percepciójának elősegítése is hozzájárul ezekhez a hatásokhoz.

Az egyes beszélők közötti eltérések elemzése alapján láthattuk, hogy az(ok) a beszélő(k), akik hajlamosabbak valamely mérési paraméter mentén a várt tendenciától eltérő ejtést vagy annak megfelelőt produkálni, a másik mérésekben is hasonló értékeket mutatnak. Azaz ez alapján azt feltételezhetjük, hogy nem a követő zöngésségnek a magánhangzóra gyakorolt hatásának létét kell egyértelműen elvetnünk, hanem a beszélői ejtési stílusok mentén kell az eredményeket értelmezni.

A magánhangzók és a mássalhangzók saját jegyeinek, illetve feltehetően azok közösen gyakorolt hatását és a beszélők közötti variabilitás további részleteit természetesen egy olyan vizsgálat segíthet megállapítani, amelyben a hangkörnyezetek és a beszélők számát is emeljük. Az eredményeket részben korlátozható hatásként meg kell jegyeznünk azt is, hogy az elemzett szavakban az első szótag nyílt, azaz a mássalhangzó másik szótaghoz tartozik, míg az utolsó szótagi magánhangzó esetében egy szótagban szerepelnek. Ez a legtöbb elemzésben hasonló volt. Emellett a mondatok szerkezetének hatását (pl. prozódia) is figyelembe kell venni a következő kutatásokban.

Irodalom

- Bárkányi Zsuzsanna – Kiss Zoltán 2009. Word-final fricative contrasts in Hungarian. A phonetic approach. Előadás a BuPhoC 2009 nov. 5-i ülésén. Letölthető: <http://budling.nytud.hu/~cash/papers/buphoc09-slide.pdf>
- Boersma, Paul – Weenink, David 2013. *Praat: doing phonetics by computer [Computer program]. Version 5.3.* <http://www.praat.org>. (Letöltés ideje: 2013. júl. 9.)
- Chen, Matthew 1970. Vowel length variation as a function of the voicing of the consonant environment. *Phonetica* 22. 129–159.
- Dunn, Margaret 1993. *The phonetics and phonology of geminate consonants*. PhD-thesis. Yale University, New Heaven.

- Fletcher, Janet 2010. The prosody of speech: Timing and rhythm. In Hardcastle, William J. – Laver, John. – Gibbon, Fiona E. (szerk.): *The handbook of phonetic sciences*. Wiley–Blackwell Publishing, Oxford. 523–602.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- Gósy Mária – Beke András 2010. Magánhangzó-időtartamok a spontán beszédben. *Magyar Nyelvőr* 134/2. 140–165.
- Grácz Tekla Etelka 2012. *Zörejhangok akusztikai fonetikai vizsgálata a zöngésségi oppozíció függvényében*. PhD-értekezés. ELTE, Budapest.
- Javkin, Hector R. 1976. *The perceptual basis of vowel duration differences associated with the voiced/voiceless distinction. Report of the Phonology Laboratory I*. University of California, Berkely. 78–92.
- Kassai Ilona 1979. *Időtartam és kvantitás a magyar nyelvben*. Nyelvtudományi Értekezések 102. Akadémiai Kiadó, Budapest.
- Kassai Ilona 1982. A magyar beszédhangok időtartamviszonyai. In Bolla Kálmán (szerk.): *Fejezetek a magyar leíró hangtanból*. Akadémiai Kiadó, Budapest. 115–154.
- Keating, Patricia A. 1984. Universal phonetics and the organization of grammars. *UCLA Working Papers of Phonetics* 55. 35–49.
- Kohári Anna 2010. *Magyar magánhangzók kontextusfüggő időviszonyai*. Szakdolgozat. ELTE, Budapest.
- Kovács Magdolna 2002. *Tendenciák és szabályszerűségek a magánhangzó-időtartamok produkciójában és percepciójában*. Debreceni Egyetem Kossuth Egyetemi Kiadója, Debrecen.
- Ladefoged, Peter – Maddieson, Ian 1996. *The sounds of the world's languages*. Blackwell, Oxford.
- Maddieson, Ian 1997. Phonetical universals. In Hardcastle, William J. – Laver, John (eds.): *The handbook of phonetic sciences*. Blackwell, Oxford. 619–639.
- Mády, Katalin – Reichel, Uwe 2007. Quantity distinction in the Hungarian vowel system – Just a theory or also reality? *Proceedings 15th International Congress of Phonetic Sciences*, Saarbrücken.
- Magdics Klára 1966. A magyar beszédhangok időtartama. *Nyelvtudományi Közlemények* 68. 125–139.
- Noteboom, Sieb 1997. The prosody of speech: Melody and rhythm. In Hardcastle, William J. – Laver, John (ed.): *The handbook of phonetic sciences*. Blackwell, Oxford. 640–673.
- Ohala, John J. 1983. The origin of the sound patterns in vocal tract constraints. In MacNeilage, Peter F. (ed.): *The production of speech*. Springer Verlag, New York. 189–216.
- Olaszy Gábor 2000. Kísérlet a magyar beszédhangok specifikus időtartamainak meghatározására folyamatos beszédre. *Beszédkutató 2000*. 26–38.
- Raphael, Lawrence J. – Dorman, Michael F. – Freeman, Francis 1975. Vowel and nasal duration as cues to voicing in word-final stop consonants: Spectrographic and perceptual studies. *Journal of Speech and Hearing Research* 18/3. 389–400.
- Van Santen, Jan P. H. 1992. Contextual effects of vowel duration. *Speech Communication* 11. 513–546.
- Stevens, Kenneth N. 1998. *Acoustic phonetics*. The MIT Press, Cambridge–Massachusetts–London.

ZÖNGÉTTLEN EXPLOZÍVÁK IDŐVISZONYAI A BESZÉDTÍPUS ÉS AZ EGYÉNI SAJÁTOSSÁGOK TÜKRÉBEN

Neuberger Tilda

Bevezetés

A beszédhangok időszerkezeti vizsgálata nagy múltat tekint vissza a szakirodalomban. Számos nemzetközi és hazai kutatás foglalkozott a felpattanó zárhangok időviszonyaival. A teljes hangidőtartam mellett a mássalhangzó egyes részeinek időtartama külön-külön is meghatározható, ugyanis az explozívaképzés a különböző artikulációs mozgások egymás utáni végrehajtásával történik, ami tükröződik az összetett akusztikai szerkezetben (Henton et al. 1992). A magyar explozíváknál a zárszakasz alatt fojtott vagy előzőnge (zöngés zárhang esetén) vagy néma fázis (zöngétlen zárhang esetén) látható az akusztikai lenyomaton. A zárfeloldástól a követő zöngé megindulásáig tartó szakasz a zöngékezdési idő (VOT, az angol „voice onset time” rövidítéséből), amely a zöngétlen zárszakasz esetén 0 ms-os vagy pozitív érték, míg zöngés zárszakasz esetén negatív értékű (Lisker–Abramson 1964; Lieberman–Blumstein 1988). Az explozívák tipikus realizációit aerodinamikai okok, koartikulációból vagy lazább artikulációból eredő hatások megváltoztathatják, így létrejöhetnek korai vagy további felpattanások is (Keating et al. 1980; Fuchs 2005; Lousada et al. 2010; Grácz–Kohári 2012). Több tanulmány megemlíti nem felpattanó zárhangként realizálódó explozívákat is, amelyek esetében felpattanás nem adatolható, így a VOT nem adható meg pontosan (Shadle 1997; Gósy–Ringen 2009; Grácz 2011).

A zárhangok objektív időtartamát vizsgáló kutatások kimutatták, hogy több tényező gyakorolhat hatást az időtartamértékekre. Ezek lehetnek nyelvi vagy nem nyelvi befolyásoló tényezők, mint például az explozíva képzési helye (Klatt 1975; Kassai 1982; Cho–Ladefoged 1999; Gósy 2000; Kovács 2001), a hangsorban elfoglalt helye és a hangsúlyviszonyok (Lisker–Abramson 1967; Fant 1973; Gósy–Ringen 2009; Gósy 2010; Grácz 2012), a hangkörnyezet vagy kontextus (Lisker–Abramson 1964; Kassai 1979; Rochet–Fei 1991; Chen et al. 2007; Grácz–Kohári 2012; Grácz 2013) vagy a beszéd típusa (Lisker–Abramson 1967; Baran et al. 1977; Krull 1991; Gósy 2000, 2001; Bóna 2011), továbbá a beszélő életkora (Menyuk–Klatt 1975; Ryalls et al. 1997; Torre–Barlow 2009; Bóna 2011; S. Tar 2013), illetve neme (Koenig 2000; Whiteside et al. 2004; Oh 2011; Neuberger–Grácz 2013). Az explozívák időszerkezetére a felsorolt tényezőkön kívül a beszélők egyéni artikulá-

ciós stratégiái, illetve anatómiai eltérései (pl. a szájüreg térfogata) is hatnak (Baran et al. 1977; Allen et al. 2003; van Alphen–Smits 2004).

A beszéd típusának (vagy stílusának) az artikulációra gyakorolt hatása régóta ismeretes a szakirodalomban. Labov (1972) szociolingvisztikai megközelítésű munkájában a beszélő beszédre fordított figyelmének mértéke alapján különböztetett meg beszédstílusokat (a spontán beszéd hétköznapi lezser és gondozott változata, valamint szöveg-, szólista- és minimálpár-olvasás). Lindblom (1990) a beszédprodukció és -percepció összefüggéseit modellező H&H (hyper and hypo speech: túl- és alulartikulált beszéd) elméletében az artikulációt a feszesség tekintetében írja le. A két legszélsőségesebb beszédstílus, a felolvasás és a spontán beszéd ennek a skálának a két végpontja felé közelít: a felolvasás a túlartikulált, míg a spontán beszéd az alulartikulált beszédhez áll közelebb. Az artikulációs pontosság a beszédhangok képzésében, így az időtartamértékekben is megmutatkozhat. A hangidőtartam-változásokat olvasott vagy ismételt beszéd, illetve izolált szavak, mondatok ejtése és spontán beszéd között tesztelték. A felpattanó zárhangok esetében a kutatások többsége a zöngekezdési idő értékeire fókuszált.

Egy angol nyelvűre vonatkozó vizsgálat (Lisker–Abramson 1967) feltárta, hogy izolált szavakban a [p, t, k] hangok VOT-értéke hosszabb, mint mondatokba ágyazott szavakban. Szignifikáns különbséget találtak a zöngekezdési időkből a társalgásban és az ismételt beszédben mért értékek között (Baran et al. 1977). Svéd nyelvűre nem volt kimutatható a beszédstílus szignifikáns hatása az olvasott szólista és a spontán beszéd között, de a VOT-értékek spontán beszédben valamivel hosszabban realizálódtak (Krull 1991). Ezzel ellentétesen (ugyanabban a kutatásban) a zárszakasz az izolált szavakban bizonyult hosszabb időtartamúnak.

Magyar nyelven Gósy (2000) öt női adatközlő beszédében elemezve a [p, t, k] zárhangokat szignifikáns különbséget igazolt az izolált szavakban és a spontán beszédben mért VOT-értékek között. Az egyének között nem volt kimutatható ilyen típusú eltérés. Az alveoláris zöngétlen explozíva esetén nem, a bilabiális és a veláris zöngétlen explozíva esetén azonban bizonyította a beszédstílus VOT-értékekre gyakorolt hatását (Gósy 2001). Az alveoláris zárhangok a spontán közlésben, míg a bilabiális és a veláris zárhangok az izolált szavakban realizálódtak hosszabb VOT-értékekkel.

Bóna (2011) fiatalok és idősek zöngekezdési idejét vetette össze olvasott és spontán beszédben, de egyik zárhang esetében sem talált szignifikáns különbséget a beszédstílus szerint. Az idősek (70–80 év közötti) beszélőknél a felolvasásból származó átlagértékek csaknem egyenlők voltak a spontán beszédben adatoltakkal, míg a fiatal (22–32 év közötti) beszélőknél a spontán beszédben kapott átlagértékek mindhárom hang esetében valamivel hosszabbak voltak. Mind a beszéd- és artikulációs tempóban, mind a zöngekezdési időben nagy egyéni különbségek jelentek meg a beszélők között. Az említett magyar vizsgálatok női adatközlők beszédére irányultak. Felmerül tehát a

kérdés, hogy hasonló mintázatot mutatnak-e az időtartamértékek a férfiak olvasott és spontán beszédében is.

A jelen kutatás célja annak vizsgálata, hogy hogyan alakulnak a zöngétlen explozívák időviszonyai a férfiak olvasott és spontán beszédében. További célunk annak kimutatása, hogy milyen jellegzetes eltéréseket eredményeznek a beszélők között az egyéni artikulációs sajátosságok. A kutatás kérdései, hogy a beszéd típus miként hat a vizsgált mássalhangzók időviszonyaira, illetve hogy egyforma mértékben, azonos paraméterekre hat-e beszélőnként. Feltételeztük, hogy (i) a teljes időtartam nagyobb mértékben függ a beszéd típusától, az explozíva belső időszerkezete pedig az egyéni ejtési sajátosságoktól, valamint hogy (ii) a beszélő sajátos artikulációja beszéd típusától függetlenül hatást gyakorol az időtartamokra.

Kísérleti személyek, anyag, módszer

Az elemzéshez tíz férfi beszélő olvasott és spontán beszédét választottuk ki a BEA adatbázisból (Gósy 2012). Életkoruk 21–29 év közötti (átlagéletkor: 24 év), mind budapesti, magyar anyanyelvű személyek. Az olvasott beszédben az adatközlők feladata 25 egyszerű és összetett mondat felolvasása volt (például *A tulipánágysok gyomlálásával üttötte el az időt, Nem kötött biztosítást ezért kisebb vagyona került a kórházi ellátás*). A spontánbeszéd felvételben munkájukról, hobbijukról beszéltek kötetlenül.

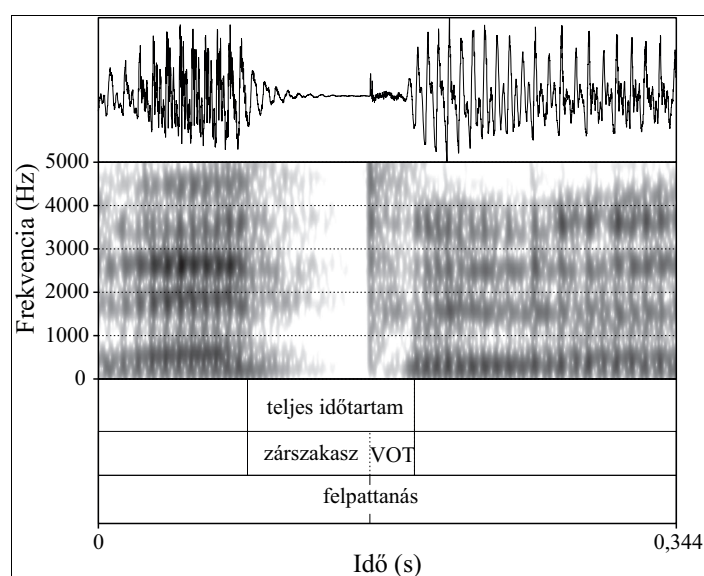
A vizsgálat anyagát összesen 286 bilabiális, 656 alveoláris és 553 veláris zöngétlen explozíva képezi. Személyenként átlagosan 150 beszédhangot elemeztünk; 14 [p], 33 [t] és 28 [k] előfordulását beszéd típusonként. Az azonos képzési helyű zárhangok körülbelül azonos arányban szerepeltek az egyes beszélők olvasott és a spontán beszédében (1. táblázat). Az összes elemzett előfordulás intervokális helyzetű volt, és szókezdő (V#CV), szó belseji (VCV) vagy szó végi (VC#V) pozíciókban fordultak elő. A hangsorbeli helyzetet tekintve mindhárom vizsgált zárhangból közel azonos arányban elemeztünk előfordulásokat a három pozíció szerint (a leggyakoribb – az adatok átlagosan 50%-a – a szó belseji pozíció volt). Jóllehet a zárhangot követő magánhangzó minősége hatással lehet a zárhangok időértékeire, főként a zöngékezdési időre (vö. Gósy 2000; Grácsi 2013), a jelen kutatásban nem e tényező mentén vizsgáltuk az explozívakat. Törekedtünk azonban az egyes képzési helyű explozívarealizációk kiegyenlítetttségére a magánhangzó-környezet szerint is, mindkét beszéd típusban.

A címkézést és a mérést a Praat 5.3 (Boersma–Weenink 2011) szoftver segítségével végeztük. Az annotálás manuálisan történt a spektrogram és az oscillogram mintázata és meghallgatás alapján. Magánhangzó-alapú szegmentálást alkalmaztunk, vagyis a vizsgált zárhang határait a megelőző, illetve a követő magánhangzó második formánsa alapján jelöltük (1. ábra). A többszöri felpattanások esetében a zárszakasz végét, illetve a VOT-értékeket az első

felpattanástól számítottuk. Figyelembe vettük az esetlegesen nem felpattanó zárhangként megvalósult előfordulások számát is.

1. táblázat: Az explozívák előfordulása a tíz beszélő olvasott és spontán beszédében

Beszélők	□□□		□□□		□□□	
	olvasott	spontán	olvasott	spontán	olvasott	spontán
1	12	13	33	36	24	27
2	11	12	32	32	25	32
3	12	20	31	35	25	25
4	13	29	31	36	27	43
5	11	16	30	34	25	35
6	12	14	32	35	24	32
7	12	17	32	38	25	34
8	11	15	32	30	25	25
9	10	13	33	37	26	21
10	12	21	32	25	25	28



1. ábra
A vizsgált fonetikai paraméterek címkézése

A mássalhangzó teljes időtartamát, a zárszakasz teljes időtartamhoz viszonyított arányát, illetve a zöngékezdesi idő értékeit elemeztük a beszédtypus és a beszélők függvényében. Az adatokat statisztikai próbáknak (Pearson-féle korrelációelemzés, egytényezős varianciaanalízis – ANOVA, UNIANOVA, Tukey-féle post hoc teszt, páros *t*-próba) vetettük alá, ehhez az SPSS 13.0 verziószámú szoftvert használtuk. A statisztikai elemzéseket 95%-os konfidenciaintervallumon végeztük el.

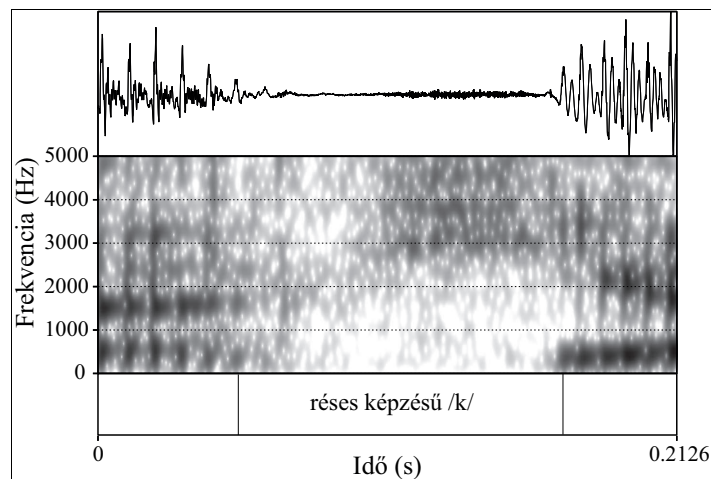
Eredmények

Az explozívák tipikus előfordulásai mellett megjelentek olyan zárhang-realizációk, amelyekben felpattanás nem volt adatolható. Az összes explozívarealizáció bizonyos hányadánál a zár feloldása egyrészt annak gyenge intenzitása miatt nem volt nyomon követhető a regisztrátumon, másrészt a beszélők réses képzéssel valószínűsítették meg az adott hangot (2. ábra). Az ilyen realizációk olvasott beszédben a /p/ esetében 9,5%-ot, a /t/ esetében 0,3%-ot, a /k/ esetében 13,1%-ot tettek ki. Spontán beszédben a /p/-nél 10,6%-os, a /t/-nél 1,2%-os, a /k/-nál pedig 12,6%-os volt az előfordulásuk. A két beszédtypusban tehát közel azonos arányban fordultak elő a tipikustól eltérő realizációs módok. Egyik zárhang esetében sem volt szignifikáns a különbség az olvasott és a spontán beszédben adatolt értékek között (páros *t*-próba: /p/ esetében: $p = 0,897$; /t/ esetében: $p = 0,192$; /k/ esetében: $p = 0,858$). Az eltérő realizációs módok olvasott és spontán beszédben megjelenő arányát beszélőnként összevetettük. A Pearson-féle korrelációelemzés szerint közepes erősségű, szignifikáns összefüggés van a két érték között, vagyis amelyik beszélőre jellemző a tipikustól eltérő artikulációs megvalósítás spontán beszédben, az a felolvasás során is gyakran így valószínűsította meg zárhangjait ($r = 0,598$; $p < 0,001$). Ezekben az esetekben – mivel a zárszakasz és a VOT pontos időtartamának meghatározása lehetetlen – csak az explozíva teljes időtartamát elemeztük.

Az explozívák teljes időtartamát tekintve különbséget feltételeztünk a beszédtypusok között. Az eredmények azonban nem mutattak jelentős különbséget az olvasott és a spontán beszédben mért átlagértékek között (2. táblázat). A három vizsgált zárhang időtartamára vonatkozó értékek a spontán beszédben azonosak vagy magasabbak voltak, mint az olvasott beszédből származó értékek. A statisztikai elemzés szerint ez a különbség csak az alveoláris explozíva esetén szignifikáns [egyváltozós ANOVA: $F(1, 655) = 6,269$; $p = 0,013$]. Az időtartamértékek szórása a [p] kivételével a spontán beszédben bizonyult nagyobbaknak, a hangok képzési helyét illetően pedig $k > t > p$ tendencia figyelhető meg az időtartamértékek szórásában mindkét beszédtypusban.

Az egyéni különbségek a zárhangok teljes időtartamában is megmutatkoztak. Mindhárom explozíva esetében szignifikáns eltéréseket találtunk az egyes beszélők között beszédtypustól függetlenül [egyváltozós ANOVA: /p/

$F(9, 266) = 2,169; p = 0,024; \eta^2 = 0,068; /t/ F(9, 637) = 4,190; p < 0,001; \eta^2 = 0,056; /k/ F(9, 533) = 3,439; p < 0,001; \eta^2 = 0,055$].



2. ábra

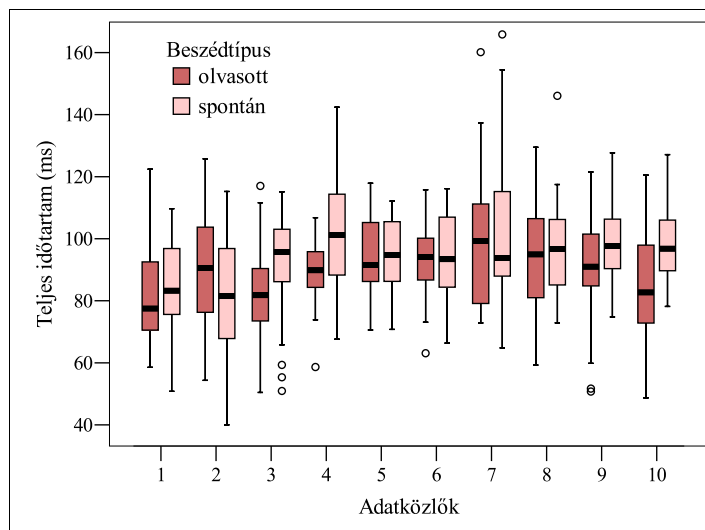
Nem tipikus explozívarealizáció

2. táblázat: Az explozívák teljes időtartamának átlaga és szórása a beszéd típus függvényében (ms)

	Olvasott	Spontán
[p]	97 (±17)	99 (±15)
[t]	90 (±17)	94 (±18)
[k]	97 (±21)	97 (±22)

Nemcsak az azonos képzési helyű hangok időtartamában különböztek jelentősen a beszélők, hanem abban is, hogy a két beszéd típusban miként alakultak az explozívák időtartamértékei. Volt olyan beszélő, akinek az olvasott beszédben mért értékei voltak magasabbak, és volt olyan is, aki a spontán beszédben valószínűleg hosszabban a hangokat. A [p] hang esetében öt beszélőnél a spontán beszédben, négyenél a felolvasásban találtunk magasabb időértékeket, egyenél (1 ms-os különbséggel) közel azonos volt a két átlagérték. Szignifikáns eltérés csak egyikükénél volt kimutatható. A [t] esetében hét beszélőnél a spontán beszédben, kettőnél az olvasott beszédbeni átlagértékek voltak magasabbak, egy beszélőnél pedig közel azonos. Három beszélőnél (4., 9., 10. számú adatközlőnél) a beszéd típus szignifikáns hatással volt az időértékekre (3. ábra). A [k] esetében négy beszélő spontán beszédbeni értékei, négy beszélő olvasott beszédbeni

értékei voltak magasabbak, kettőnél pedig kvázi egyenlő átlagértékeket találtunk. A beszéd típusok szerinti különbség csak egyiküknél volt szignifikáns.



3. ábra

A [t] teljes időtartama a tíz adatközlő olvasott és spontán beszédében

A zönggekezdési idejük alapján a három vizsgált explozíva szignifikánsan elkülönült egymástól (olvasott beszédben: $F(2, 639) = 106,604$; $p < 0,001$; spontán beszédben: $F(2, 749) = 76,708$; $p < 0,001$). A post hoc tesztek azt mutatták, hogy olvasott beszédben a veláris zárhang zönggekezdési ideje tér el szignifikánsan a másik kettőtől ($p < 0,001$ mindkét esetben), spontán beszédben pedig mindhárom zárhang VOT-je szignifikánsan eltért a másik kettőtől ($p < 0,001$ minden esetben). A korábbi szakirodalom (pl. Gósy 2001; Kovács 2001) eredményeinek megfelelően a képzési hely tükrében a legrövidebb VOT-értékkel a bilabiális, a leghosszabbal pedig a veláris explozívák realizálódtak (3. táblázat). A szórás beszéd típustól függetlenül a [k] zárhang esetében volt a legnagyobb, és a [t] esetében a legkisebb.

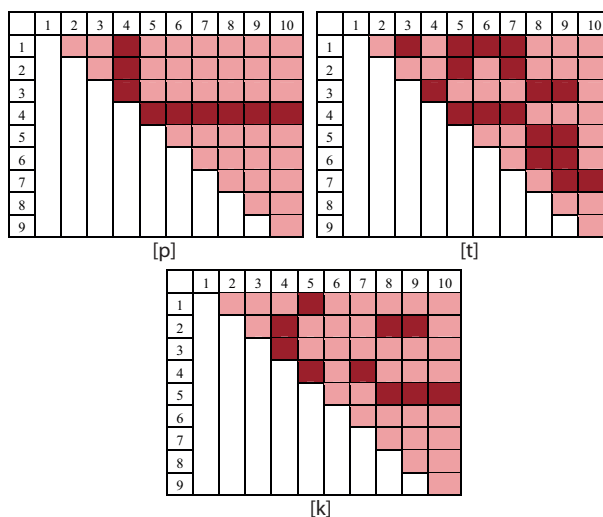
Bár a VOT átlagértékei mindhárom zárhang esetén a spontán közlésben voltak magasabbak, a statisztikai elemzés szerint a beszéd típus hatása csak a [t] esetén volt matematikailag igazolható [$F(1, 632) = 24,124$; $p < 0,001$; $\eta^2 = 0,037$].

A varianciaanalízis az egyes beszélők között szignifikáns eltérést mutatott a VOT-értékek tekintetében is, a beszéd típustól függetlenül (Univariate ANOVA): /p/ $F(9, 237) = 6,825$; $p < 0,001$; $\eta^2 = 0,206$; /t/ $F(9, 632) = 18,327$; $p < 0,001$; $\eta^2 = 0,207$; /k/ $F(9, 462) = 8,648$; $p < 0,001$; $\eta^2 = 0,144$). A Tukey-féle

post hoc teszt alapján a /p/-nél a 4. számú beszélő értékei tértek el szignifikánsan az összes többiétől, a másik két explozíva esetén változatosabb volt a beszélők közötti eltérés (4. ábra). A beszélők közötti szignifikáns eltérések a [t] esetén voltak a legkiterjedtebbek, vagyis a három zárhang közül az alveoláris VOT-értékeit tekintve különbözött egymástól a legtöbb beszélő ejtése.

3. táblázat: Az explozívák VOT-értékeinek átlaga és szórása a beszéd típus függvényében

	Olvasott (ms)	Spontán (ms)
[p]	20 (± 14)	22 (± 10)
[t]	24 (± 9)	27 (± 11)
[k]	32 (± 17)	36 (± 14)



4. ábra

A VOT-értékekre vonatkozó beszélők közötti különbségek mátrixos ábrázolása (sötét szürke = szignifikáns, világos szürke = nem szignifikáns)

A beszélők zárhangjainak VOT-értékeire a beszéd típus kismértékű hatással volt; csupán négy beszélő egy-egy zárhangjának VOT-értékei különböztek egymástól szignifikánsan a beszéd típus szerint. A beszéd típus VOT-értékekre gyakorolt hatása hármuknál a [t] hangot érintette.

A zárszakasz időtartamát az explozíva teljes időtartamához viszonyított arányában elemeztük. A VOT-értékekhez hasonlóan, a zárszakasz aránya alapján is szignifikánsan különbözött a három vizsgált explozíva (olvasott beszédben: $F(2, 639) = 107,995$; $p < 0,001$; spontán beszédben: $F(2, 749) =$

93,182; $p < 0,001$). Az egyes beszédhangok összehasonlításában a post hoc tesztek is minden esetben szignifikáns eltéréseket mutattak ($p < 0,001$). Az átlagértékeket tekintve a zárarány a legmagasabb a [p], a legalacsonyabb a [k] esetében volt (4. táblázat). Kovács (2001) vizsgálatában hasonló tendenciáról számol be, öt női adatközlő mondatfelolvasásában a zárarányok átlaga a következőképpen alakult: [p] 83%, [t] 78,5%, [k] 69%.

4. táblázat: A zárarány átlaga és szórása a beszéd típus függvényében

	Olvasott (%)	Spontán (%)
[p]	78 (± 10)	78 (± 9)
[t]	73 (± 10)	71 (± 11)
[k]	63 (± 10)	64 (± 11)

A beszéd típus szignifikáns hatással volt az alveoláris explozív zárarányára [$F(1, 649) = 7,171$; $p = 0,008$; $\eta^2 = 0,011$]; olvasott beszédben a [t] hangok szignifikánsan hosszabb zárszakasszal realizálódtak. A másik két zárhang esetében nem volt kimutatható ilyen hatás.

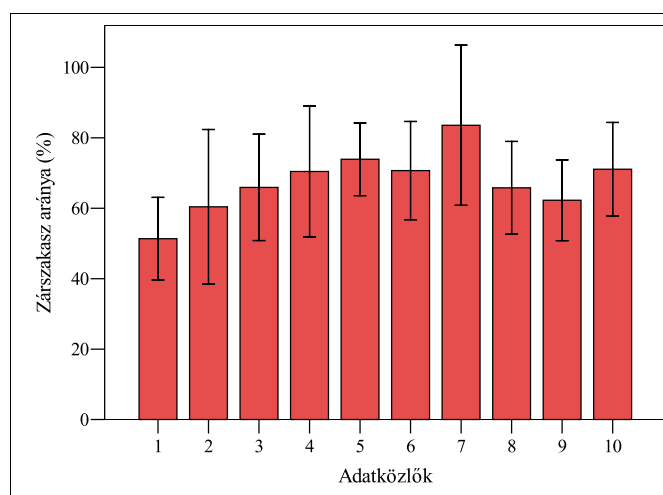
A tíz beszélő között szignifikáns eltérések mutatkoztak ebben a paraméterben is mindhárom explozív esetén beszéd típustól függetlenül [egytényezős ANOVA: /p/ $F(9, 237) = 7,691$; $p < 0,001$; $\eta^2 = 0,226$; /t/ $F(9, 632) = 26,089$; $p < 0,001$; $\eta^2 = 0,271$; /k/ $F(9, 462) = 7,790$; $p < 0,001$; $\eta^2 = 0,132$].

A zárszakaszok időtartamának arányában az adatközlők között nagy egyéni különbségeket fedezhettünk fel. Szemléltetésképpen a spontán beszédben megvalósult [t] hangok zárarányát mutatjuk be beszélőnként (5. ábra). A zárszakaszok időtartamának átlagos aránya a beszélők között 50–80%-os volt.

Végül megvizsgáltuk a beszéd típus hatását a három vizsgált paraméteren beszélőnként is. A statisztikai elemzés a legnagyobb mértékben a három zárhang közül a [t] esetében mutatta ki a beszéd típus időtartamértékekre gyakorolt hatását. Volt olyan beszélő (9. számú), akinél mindhárom paraméterre (teljes időtartam, VOT-érték, zárszakasz aránya) szignifikáns hatással volt a beszéd típus, és akadt olyan is, akinél csak egy vagy két paraméterre (6. ábra). A három időtartam közül a mássalhangzó teljes időtartama reagált a legmarkánsabban a beszéd típus hatására.

Következtetések

A jelen kutatásban a zöngétlen felpattanó zárhangok időviszonyait elemeztük tíz férfi beszélő olvasott és spontán beszédében. A vizsgált fonetikai jegek között a mássalhangzó teljes időtartama, a zárszakasz időtartamának a teljes időtartamhoz viszonyított aránya és a VOT-értékek szerepeltek. A beszéd típus és az egyéni ejtési sajátosságok hatását kívántuk kimutatni.



5. ábra

A [t] zárszakasz-időtartamának a teljes időtartamhoz viszonyított aránya beszélőnként (spontán beszéd)

a: teljes időtartam (ms)				b: VOT-érték (ms)				c: zárszakasz aránya (%)			
	a	b	c		a	b	c		a	b	c
1				1				1			
2				2				2			
3				3				3			
4				4				4			
5				5				5			
6				6				6			
7				7				7			
8				8				8			
9				9				9			
10				10				10			

[p]

[t]

[k]

6. ábra

A beszéd típus hatása a vizsgált paraméterekre beszélőnként (sötét = szignifikáns, világos = nem szignifikáns)

Mindkét beszéd típusban adatoltunk az explozívák alaprealizációjától eltérő előfordulásokat is. A tipikus zárszakasszal és felpattanással jellemezhető zárhang-előfordulások mellett felpattanás nélküli, illetőleg réshangként képzett

realizációk is megjelentek. Ezek aránya hozzávetőlegesen megegyezett olvasott és spontán beszédben, illetőleg szignifikánsan összefüggött a két beszéd-típusban. Ebből arra következtethetünk, hogy az ilyen artikulációs megvalósítás a beszélők sajátos ejtésére jellemző, és nem a spontán beszédnek (a felolvasáshoz viszonyított) kevésbé gondozott voltából adódik.

Hipotézisünk volt, hogy a beszéd-típus és az egyéni ejtési sajátosságok különbözőképpen hatnak az explozívák teljes időtartamára, illetve belső időszervezetükre. Eredményeink szerint a beszéd-típus hatása eltérő hatásfokkal jelent meg a különböző időparamétereken. A három vizsgált paraméter közül a legnagyobb hatással a zárhangok teljes időtartamára volt. A VOT-értékeknek a $k > t > p$ tendenciáját több nyelvre is igazolták már (angol: Lisker–Abramson 1967; svéd: Fant 1973; magyar: Gósy 2001; Kovács 2001), és a jelen kutatás eredményei is megerősítik ezt. Eredményeink arra is fényt derítettek, hogy ez a tendencia beszéd-típustól független, ugyanis olvasott és spontán beszédben is hasonlóképpen valósult meg.

A három, eltérő képzési helyű zöngétlen explozíva közül az alveoláris időértékeket tekintve szignifikáns különbséget adatoltunk az olvasott és a spontán beszéd között, míg a bilabiális és a veláris zárhangok esetében nem. Az eredmények azt is kimutatták, hogy a beszéd-típus nem minden adatközlőnél hat azonos módon a vizsgált hangok különböző időparamétereire.

Másik hipotézisünk – miszerint a beszélő sajátos artikulációja beszéd-típustól függetlenül hatást gyakorol az időtartamokra – beigazolódott. A beszélők közötti variabilitás erőteljesebben éreztette hatását, mint a beszéd-típus. Erre utal az, hogy az egyes adatközlők időtartamértékei az explozívák mindhárom vizsgált paraméterében szignifikáns különbségeket mutattak beszéd-típustól függetlenül. A kutatás eredményeiből tehát az látszik, hogy a beszélő mint faktor jobban befolyásolja az időtartamértékeket, mint a beszéd típusa.

A különböző beszéd-típusokban megjelenő egyéni ejtési sajátosságok vizsgálata az alkalmazott fonetika több területén is felhasználható, valamint alkalmazható lehet a logopédiai vagy a beszédtechnikai, retorikai gyakorlatban.

Irodalom

- Allen, J. Sean – Miller, Joanne L. – DeSteno, David 2003. Individual talker differences in voice-onset-time. *Journal of Acoustic Society of America* 113/1. 544–552.
- Van Alphen, Petra M. – Smits, Röel. 2004. Acoustical and perceptual analysis of the voicing distinction in Dutch initial plosives: the role of prevoicing. *Journal of Phonetics* 32/4. 455–449.
- Baran, Jane A. – Laufer, Marsha Z. – Daniloff, Ray 1977. Phonological contrastivity in conversation: A comparative study of voice onset time. *Journal of Phonetics* 5/4. 339–350.
- Boersma, Paul – Weenink, David 2013. *Praat: doing phonetics by computer [Computer program]. Version 5.3.* <http://www.praat.org>. (Letöltés ideje: 2013. júl. 9.)

- Bóna Judit 2011. A [p, t, k] mássalhangzók zöngékezdési ideje idők és fiatalok spontán beszédében és felolvasásában. *Beszéd kutatás 2011*. 61–72.
- Chen, Li-Mei – Chao, Kuan-Yi – Peng, Jui-Feng. 2007. VOT productions of word-initial stops in Mandarin and English: A cross-language study. In *Proceedings of the 19th Conference on Computational Linguistics and Speech Processing*. 303–317. <http://www.newdesign.aclweb.org/anthology-new/O/O07/O07-2004.pdf>
- Cho, Taehong – Ladefoged, Peter. 1999. Variations and universals in VOT: Evidence from 18 languages. *Journal of Phonetics* 27. 207–229.
- Fant, Gunnar 1973. *Speech sounds and features*. MIT Press, Cambridge, MA.
- Fuchs, Susanne 2005. *Articulatory correlates of the voicing contrast in alveolar obstruent production in German*. Doctoral dissertation. Queen Margaret University College. http://193.62.47.80/casl/pubs/Fuchs_2005_PhD_ZAS_version.pdf.
- Gósy Mária 2000. A [p, t, k] mássalhangzók zöngékezdési ideje. *Magyar Nyelvőr* 124. 195–204.
- Gósy, Mária 2001. The voice onset time of the Hungarian voiceless plosives in words and in spontaneous speech. *International Journal of Speech Technology* 3–4. 155–164.
- Gósy, Mária 2010. Phonetic variation in Hungarian /t/. *Studia Slavica Hungarica* 55. 255–261.
- Gósy Mária 2012. Multifunkcionális beszélt nyelvi adatbázis – BEA. In Prószyk Gábor – Várad Tamás (szerk.): *Általános Nyelvészeti Tanulmányok XXIV. Nyelvtchnológiai kutatások*. Akadémiai Kiadó, Budapest. 329–349.
- Gósy, Mária – Ringen, Catherine O. 2009. *Everything you always wanted to know about VOT in Hungarian*. Elhangzott: International Conference on the Structure of Hungarian 2009. Budapest, 2009. szeptember 1. http://ics9.unideb.hu/pph/handout/Ringen_Gosy_handout.pdf.
- Grácsi Tekla Etelka 2011. Intervokális explozívák a zöngésségi oppozíció függvényében. *Beszéd kutatás 2011*. 46–60.
- Grácsi Tekla Etelka 2012. Az explozívák időszerkezete spontán beszédben. In Gósy Mária (szerk.): *Beszéd, adatbázis, kutatások*. Akadémiai Kiadó, Budapest. 67–90.
- Grácsi Tekla Etelka 2013. Explozívák és affrikáták zöngésségének időviszonyai. *Beszéd kutatás 2013*. 94–120.
- Grácsi Tekla Etelka – Kohári Anna 2012. A zöngékezdési idő egy módszertani kérdés függvényében. In Markó Alexandra (szerk.): *Beszédtudomány. Az anyanyelv-elsajátítástól a zöngékezdési időig*. ELTE Bölcsészettudományi Kar – MTA Nyelvtudományi Intézet, Budapest, 228–248.
- Henton, Caroline – Ladefoged, Peter – Maddieson, Ian. 1992. Stops in the world's languages. *Phonetica* 49. 65–101.
- Kassai Ilona 1979. *Időtartam és kvantitás a magyar nyelvben*. Nyelvtudományi Értekezések 102. Akadémiai Kiadó, Budapest.
- Kassai Ilona 1982. A magyar beszédhangok időtartamviszonyai. In Bolla Kálmán (szerk.): *Fejezetek a magyar leíró hangtanból*. Akadémiai Kiadó, Budapest, 115–154.
- Keating, Patricia – Westbury, John R. – Stevens, Kenneth N. 1980. Mechanisms of stop-consonant release for different places of articulation. Elhangzott: *The Spring meeting of the Acoustical Society of America* 1980. http://www.linguistics.ucla.edu/people/keating/KeatingWestburyStevens_ASA1980.pdf

- Klatt, Dennis H. 1975. Voice onset time, frication, and aspiration in word-initial consonant clusters. *Journal of Speech and Hearing Research* 18. 686–706.
- Koenig, Laura L. 2000. Laryngeal factors in voiceless consonant production in men, women, and 5-year olds. *Journal Speech, Language and Hearing Research* 43. 1211–1228.
- Kovács Magdolna 2001. A zöngétlen zárhangok és affrikáták időszervezete. *Beszéd-kutatás* 2001. 46–60.
- Krull, Diana 1991. VOT in spontaneous speech and in citation form words. *PERILUS XII*. 101–107.
- Labov, William 1972. *Sociolinguistic patterns*. University of Pennsylvania Press, Philadelphia.
- Lieberman, Philip – Blumstein, Sheila E. 1988. *Speech physiology, speech perception, and acoustic phonetics*. Cambridge University Press, Cambridge.
- Lindblom, Björn 1990. Explaining phonetic variation: A sketch of the H&H theory. In Hardcastle, William J. – Marchal, Alain (eds.): *Speech production and speech modeling*. Kluwer, Dordrecht. 403–440.
- Lisker, Leigh – Abramson, Arthur S. 1964. A cross-language study of voicing in initial stops: Acoustical measurements. *Word* 20. 384–422.
- Lisker, Leigh – Abramson, Arthur S. 1967. Some effects of context on voice onset time in English stops. *Language and Speech* 10/1. 1–28.
- Lousada, Marisa – Jesus, Luis M. T. – Hall, Andreia 2010. Temporal acoustic correlates of the voicing contrast in European Portuguese stops. *Journal of the International Phonetic Association* 40/3. 261–276.
- Menyuk, Paula – Klatt, Mary 1975. Voice onset time in consonant cluster production by children and adults. *Journal of Child Language* 2/2. 223–231.
- Neuberger Tilda – Grácsi Tekla Etelka 2013. Az alveoláris zöngétlen explozívá variabilitása. *Beszéd-kutatás* 2013. 160–172.
- Oh, Eunjin 2011. Effects of speaker gender on voice onset time in Korean stops. *Journal of Phonetics* 39. 59–67.
- Rochet, B. L. – Fei, Yanmei 1991. Effect of consonant and vowel context on Mandarin Chinese VOT: Production and perception. *Canadian Acoustics* 19/4. 105–106.
- Ryalls, John – Cliche, Annie – Fortier-Blanc, Kulie – Coulombe, Isabelle – Prud'Hommeaux, Annie 1997. Voice-onset time in younger and older French-speaking Canadians. *Clinical Linguistics and Phonetics* 11/3. 205–212.
- Shadle, Christine H. 1997. The aerodynamics of speech. In Hardcastle, William J. – Laver, John (eds.): *The handbook of phonetic sciences*. Blackwell, Oxford. 33–64.
- S. Tar Éva 2013. A zöngésségi kontraszt elsajátításának mintázata atipikus nyelvfejlődés esetén. *Beszéd-kutatás* 2013. 194–209.
- Torre III, Peter – Barlow, Jessica A. 2009. Age-related changes in acoustic characteristics of adult speech. *Journal of Communication Disorders* 42/5. 324–333.
- Whiteside, Sandra P. – Henry, Luisa – Dobbin, Rachel. 2004. Sex differences in voice onset time: A developmental study of phonetic context effects in British English. *Journal of the Acoustical Society of America* 116/2. 1179–1183.

A kutatást a 108762 számú OTKA-pályázat támogatta.

AZ IRREGULÁRIS ZÖNGEMINŐSÉG GYAKORISÁGA ÉS POZÍCIÓI KÜLÖNFÉLE SPONTÁN BESZÉDHELYZETEK BEN

Markó Alexandra

Bevezetés

A zöngé egyszerű meghatározással a hangszalagok rezgésének eredményeképpen létrejövő hang. A beszéd alapját képező zöngé természetes módon közvetít információkat a beszélőről: a zöngé minősége, jellegzetességei alapján következtetünk egyebek mellett a beszélő korára, nemére, fizikai és pszichés állapotára, akár még társadalmi státusára is (Laver 1980), szokásaira, életvitelére (pl. dohányzás).

A (modális) zöngét a szakirodalom a hangszalagok kváziperiodikus rezgéseként határozza meg (vö. pl. Gósy 2004). A zöngéképzés egyes esetekben azonban (szándékosan vagy a beszélő akaratától függetlenül) eltérhet ettől, és a fonáció irregulárisba (glottalizáció, creaky voice) válthat. A glottalizáció jelenségét sokáig pusztán a patológiás zöngé sajátosságának tartották, mígnem – nagyjából az 1960-as évektől kezdődően – a szerzők szembesültek azzal a ténnyel, hogy az egészséges/tipikus beszélők produkciójában sem ritka ez a forma (Gerratt–Kreiman 2001). Azóta a világ számos nyelvében többféle funkcióját tárták fel, kidolgozták meghatározásának módszertanait, dokumentálták megjelenési módozatait, illetve kísérletet tettek arra is a kutatók, hogy a patológiás és az ép beszéd különbségeit meghatározzák a glottalizáció vonatkozásában.

Az elnevezését tekintve az angol nyelvű szakirodalomban jó néhány, egymást részben vagy egészben átfedő terminust találunk a jelenségre: *vocal fry*, *glottal fry*, *pulse phonation*, *creak*, *creaky voice*, *laryngealization*, *glottalization*. A magyarban sem egységes a terminológia, találkozunk az angolból átvett *glottalizáció* megnevezéssel, illetőleg olyan jelzős szerkezetekkel, mint *érdes*, *rekedtes*, *nyikorgó zöngé*, *laringalizált*, *csikorgó beszéd* (a terminusok áttekintését lásd Böhm–Ujváry 2008).

A zöngeminőség irreguláris minősítése nem mindig egyértelmű, hiszen az irregularitás is többféle formában realizálódhat. Huber (1988) disszertációjában négy kategóriába sorolta a megvalósulásokat: 1. „recsegő” zöngé (creaky voice), azaz irregularitás periódusról periódusra; 2. „recsegés” (creak), azaz hosszan tartó alacsony alapfrekvencia az egyes zöngperiódusok csillapításával kísérve; 3. diplofónia, azaz az egymást követő periódusok alakjában, amplitúdójában vagy periódusidejében megjelenő változatosság; 4. gégezá-

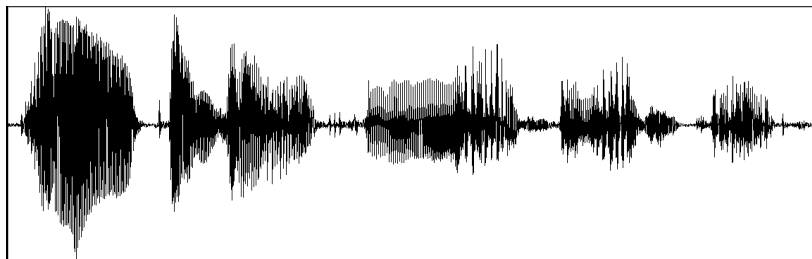
hang, azaz egyetlen periódust érintő irregularitás (idézi Redi–Shattuck–Hufnagel 2001). Batliner és munkatársai (1993) például négy beszélő mintegy 30 percnyi spontán és olvasott anyagában hatféle (5 kategória + 1 vegyes) irreguláris zöngétípust különítettek el, hat szempont alapján. Dilley és munkatársai (1996) 5 beszélő felolvasásában, rádióhírekben vizsgálták a szó eleji magánhangzókat az irreguláris zöngeminőség szempontjából. Kutatásukban négyféle realizációt határoztak meg. Az irregularitás leggyakrabban egymást szabálytalan időközökben követő periódusokként jelent meg, amelyek gégezárhanggal is kombinálódhattak. Az alaphangmagasság hirtelen le- és fellépése a gégezárhang hatását kelti a percepcióban, de az f_0 hirtelen csökkenése vagy növekedése önmagában szintén glottalizációnak minősült. Végül Houde és Hillenbrand (1994) alapján „virtuális glottalizáció”-nak nevezték azokat a megvalósulásokat, amelyekben az amplitúdó csökkenése és növekedése okozza a glottalizációra jellemző perceptuális élményt.

Az egyes nyelvekben a glottalizációnak többféle funkcióját adatolták, és természetesen egy adott nyelvben egyszerre több szerepben is megjelenhet ez a zöngeminőség, akár pozíciótól függően is. A glottalizációnak nem minden funkciója nyelvi természetű, és viszonylag kevés nyelvben tulajdonítanak neki fonológiai szerepet. Ugyanakkor igen sok kutatás igazolta egyéb beszédbe- (kommunikációs és más) funkcióit, amelyek részben beszédfiziológiai, részben szociokulturális okokkal függnek össze. A következőkben áttekintjük, hogy a glottalizációnak milyen funkcióit tárták fel eddig a magyarban.

Már a *Nyelvművelő kézikönyv*ben olvashatunk arról, hogy szakasz/mondatvégen glottalizáció jelentkezhetsz (Elekfi 1983: 774). Kísérletekkel is igazolást nyert, hogy a glottalizáció gyakran jelentkezik mondat- vagy közlésvégen – mind olvasott, mind spontán beszédben (Böhm–Ujváry 2008; Markó 2009, 2010, 2011). Felolvasásban a magyar anyanyelvű beszélők elsősorban a mondatvég (a mondathatárhoz illeszkedő frázishatár) közeledtén a jelzésére használják a glottalizációt, a tagmondathatárok esetében nem vagy alig. A glottalizációs határjelzés kifejezetten a szerkezet/frázis végén (ahhoz közeledve, az utolsó néhány szótagon) jelenik meg (vö. 1. ábra), és olykor áthúzódik a következő frázis elejére.

Szóhatáron, két magánhangzó találkozásánál a magyarban is gyakori a glottalizáció (akárcsak az angolban): a V#V kapcsolatok mintegy harmadára jellemző az irreguláris zöngképzés (felolvasott mondatok esetében). A szó eleji glottalizáció (V#V kapcsolatokban) akkor a legjellemzőbb, ha a szóhatár (magánhangzóra végződő és magánhangzóval kezdődő szavak találkozása) intonációs frázis határára esik, ekkor 50–80%-ra nő a jelenség előfordulási gyakorisága (Markó 2011, 2012a). Előfordul irreguláris zöngképzés olyan esetekben is, amikor az egymást követő szavak egy prozódiai egységhez tartoznak. Ekkor feltehetően a gondos ejtés szándéka vezeti a beszélőt (vö. Elekfi 1992) – az a célja, hogy elkerülje az olyan ejtési automatizmusokat, mint a hiátustöltés vagy a törlés (egybeolvadás). Például az *engedi át*, a *belga*

író, a *szemközti apartmanban* kapcsolatokban a hiátustöltéssel történő realizációt; a *jó oktató*, *helyreállító operáció* stb. jelzős szerkezetek esetében a törlést helyettesítheti a glottalizáció.



[ɔ m ε j ε k ε d ʃ ɔ n u : t l ɔ n u l ɛ l f o ʃ ɔ s t ɒ ŋ k]

1. ábra

Glottalizált mondatvég (*amelyeket gyanútlanul elfogyasztunk*) hullámformája (BEA adatbázis, olvasott szöveg, fiatal női beszélő) – az utolsó néhány szótagon jól látszik az irregularitás (a transzkripcióban [ɒ] jelöli)

Az irreguláris zöng gyakorisága alapvetően egyénfüggő, de befolyásolhatja a beszédmód, a nem, az életkor és más tényezők is. Felolvasott szövegek elemzése alapján megállapították, hogy a fiatal és középkorú magyar beszélők (nemtől függetlenül) mintegy a szótagok 10–12%-át ejtik glottalizáltan (Böhm–Ujváry 2008: 10 férfi, 2 női beszélő; Markó 2011: 10 női beszélő) ebben a beszédmódban. Idős (70 év feletti) női beszélőknél ez az arány 14%, tehát nincs jelentős eltérés a korosztályok között (Markó–Bóna 2012: 7 beszélő). Spontán beszédben összevetve fiatal és idős nők glottalizációs gyakoriságát Markó és Bóna (2012) azt találta, hogy a fiatalok (20–32 évesek) spontán beszédében mintegy 14% a glottalizált szótagok aránya, az idősekében 16,5%. Az időseknél tehát mind az olvasott, mind a spontán beszédben gyakoribb a glottalizáció, de az eltérés nem jelentős.

Az eddigi legnagyobb anyagon (30 beszélő spontán és olvasott beszédén) és a legtöbbféle korcsoportban, mindkét nem körében végzett vizsgálatban az átlagos előfordulás 21,3% volt az olvasott, 25,6% a spontán beszédben (Markó 2012b). A legnagyobb arányban a középkorú nők interjúiban jelent meg a glottalizáció, itt átlagosan a szótagok 34,2%-a volt ilyen (szórás: 7,6%). A legkevesebb glottalizált szótag a fiatal férfiak felolvasásában volt adatolható, 14,7% (szórás: 2,8%). A legritkábban glottalizáló beszélő egy idősebb férfi volt, az ő felolvasásában a szótagok 4,9%-a, spontán beszédében 6,0%-a valósult meg részben vagy egészben irreguláris zöngével. Az egy középkorú nő spontán beszédében adódott 47,4%-os gyakorisági érték volt a teljes vizsgálati csoportban mért maximum.

A beszélők közötti variancia nemcsak a glottalizáció gyakoriságában jelentkezik, hanem abban is, hogy milyen funkciókban alkalmazzák ezt a zöngéképzési módot. Böhm és Ujváry (2008) kimutatta, hogy a mondat végi helyzetben a glottalizációs gyakoriság az egyes beszélőkre jellemző tulajdonság (egyazon beszélő glottalizációs gyakorisága nem változott meg ugyanazon szöveg többszöri felolvasása esetén sem). A mondathatár glottalizációval történő megjelölése még azoknál a beszélőknél is relatíve gyakori, akiknek a zöngéképzése egyébként ritkán vált át irregulárisba. Egy kutatás alapján minél kevesebbet glottalizál egy beszélő, annál valószínűbb, hogy ezt mondatzáró frázishatáron teszi (Markó 2011). A magánhangzók határának jelölését tekintve is eltérő mintázatokat látunk a beszélők között: van olyan adatközlő, aki 30-szor olyan gyakran alkalmazza az irreguláris zöngét az egymást követő magánhangzók elkülönítésére, mint más célra; egy másik beszélő zöngéképzését ugyanakkor nagymértékben jellemzi az irregularitás, de a magánhangzók között nagyon ritkán alkalmazza ezt a zöngéképzési módot (Markó 2012a).

A beszélők abban a tekintetben is különböznek a glottalizációs gyakoriságot illetően, hogy mennyire tudatos hanghasználok. Egy kutatás szerint (Markó 2013) a tipikus beszélők a közlés lezárásának jelzésére gyakran alkalmaznak glottalizációt, míg ez a hivatásos beszélőkre (rádiós hírolvasókra) alig volt jellemző. Minden, az utóbbi csoportba tartozó adatközlő esetében a magánhangzók határának jelölése volt az irreguláris zöngé elsődleges funkciója. Mindez arra enged következtetni, hogy a megnyilatkozás végi glottalizációt (Elekfi [1983] szavaival a „nyekergésfélébe” hajló lezárást) kerülik a média beszédtechnikai szempontból tudatosabb beszélői.

Gósy és munkatársai pszicholingvisztikai szempontból a következőképpen határozzák meg a spontán beszédet: „A beszélő spontán módon beszél akkor, ha nem tudja (nem akarja, nem áll módjában) közlendőjét részletesen át gondolni, a meghangosítást tartalmi és formai szempontból pontosan megtervezni (szemantikai relációk, grammatikai szerkezetek, szövegszerűség, lexikális válogatás, szegmentális és szuprasegmentális artikulációs kontroll stb.)” (2009: 170). A kutatási eredmények arra hívják fel a figyelmet, hogy a spontán beszéd nem homogén kategória, hiszen ugyanazon személyek kvázimonologikus (interjú) és dialogikus spontán beszédének szuprasegmentális szerkesztettsége jellegzetesen eltér egymástól. A beszédszakaszok időtartama a társalgásban mintegy 10%-kal rövidebb, ugyanennyivel gyorsabb az artikulációs tempó. Az interjú szünetidőtartamai szignifikánsan hosszabbak voltak, mint a társalgáséi, ugyanakkor a társalgásban 40%-kal gyakrabban fordul elő szünet mondategységhatáron, azaz tervezettebb a szünethasználat. A társalgás jóval nagyobb arányban tartalmaz – dallammenettel is – lezárt közléseket, mint a kvázimonologikus szövegek (Markó 2005). A hezitálások ritkábbak és rövidebbek a társalgási, mint az interjúhelyzetben (Bóna 2013). Bár mind az interjú, mind a társalgás spontán beszédmódot indukál a szó pszicho-

lingvisztikai értelmében, a társalgásban előre tervezettedben tud megnyilvánulni a beszélő, mint az interjúhelyzetben, ahol folyamatos beszédprodukcióna kényszerül. Ezzel szemben a társalgásban mások fordulói alatt a tervezési folyamatok részben lezajlanak, a percepciók működése révén is már előhívódnak a mentális lexikonból egységek, szerkezetek stb.

Mindezek alapján felmerül a kérdés, hogy az interjú- és a társalgáshelyzet között megfigyelhető-e eltérés a glottalizációs sajátosságokat illetően is. Ennek megfelelően a jelen kutatásban az irreguláris zöng gyakoriságát és pozícióit, illetve funkcióit elemeztem ugyanazon személyek interjú- és társalgáshelyzetben rögzített beszédprodukcójában. Hipotézisem szerint az irreguláris zöngeminőség ritkább a társalgásban, ugyanakkor gyakrabban tölt be határjelző szerepet, mint az interjúhelyzetben.

Kísérleti személyek, anyag, módszer

A kutatáshoz a Beszélt nyelvi adatbázis (BEA, l. Gósy et al. 2012) 10 adatközlőjének interjút és társalgásfelvételét használtam fel. A beszélők között a nemek aránya kiegyenlített volt, életkoruk 20 és 60 év közé esett, az átlagéletkor 40,3 év. Az interjúkból elemzett beszédrészek (csak az adatközlő beszédét tekintve) összesen mintegy 48 percet tettek ki, a szótagszám ebben az alkorpuszban 10 887. A társalgásokból összesen 39 percet elemeztem (pusztán az adatközlők beszédét véve számításba), amely alatt összesen 9442 szótag hangzott el.

A címkézés a Praat 5.3 szoftverrel (Boersma–Weenink 2013) történt, az elhangzott szöveget a beszédszakaszok (két szünet között megvalósuló szöveges egység) szintjén lejegyeztem, majd címkéztem az (akár részben, akár egészben) glottalizált szótagokat. Jelöltem a szótag pozícióját a beszédszakaszban, az elemi mondatban,¹ az intonációs frázisban és a szóban. A szöveg szótagokon túlmenően a kitöltött szünetek (hezitálások) esetében is jelöltem, ha – akár részben, akár egészben – irreguláris zöngével valósultak meg.

A glottalizált realizációk elemzése a korábbi kutatások módszertanához igazodva (pl. Dilley et al. 1996; Böhm–Ujváry 2008) vizuális és auditív információk együttes figyelembevételével történt. A Praatban megjelenítettem a hullámformát és a hangszínképet, szükség esetén a dallam- és az intenzitás-görbét is; illetőleg folyamatosan és többször ellenőrizve hallgattam le a hanganyagokat. Akusztikai szempontból glottalizáltak minősítettem az adott szövegrészt, ha az alapperiódusok időtartama vagy amplitúdója hirtelen jelentősen megváltozott (idetartozónak vettem a hangszalagzár/gégezárhang jelenségét is, vö. pl. Dilley et al. 1996). Percepciók szempontként tekintetbe vettem, hogy a szegmentum hangszínezete jól hallhatóan érdes, rekedtes.

¹ „Olyan szemantikai, fonológiai és szintaktikai szerkezet, amely jelentésében egy jelenetet mutat be, és a környezetében lévő hasonló szerkezetekhez nyelvi módokon kapcsolódik” (Tolcsvai Nagy 2011: 124).

Elemeztem és a beszédhelyzetek (interjú vs. társalgás) tekintetében összehasonlítottam a glottalizált szótagok arányát és a glottalizált szótagszekvenciák hosszát. Megvizsgáltam a glottalizált szekvenciák helyzetét a szavakon, az intonációs frázisokon, a beszédszakaszokon és az elemi mondatokon belül. A második hipotézis vizsgálatához csoportosítva is elemeztem a határpozíciójút (az egységek elején, végén és határán) és az ettől eltérő előfordulásokat. Ezzel kapcsolatosan meg kell jegyezni, hogy bár az egységek elején és végén jelentkező glottalizációt határpozíciójúnak tarthatjuk, a beszédszakaszok esetében valódi határon megjelenő, vagyis az egyik egység végéről a következő elejére átívelő glottalizációt nem adatoltam. A beszédszakaszok között ugyanis per definitionem szünet jelentkezik, és a néma szünet, azaz a hangképzés felfüggesztése hatással van a zöngeminőségre.

Megvizsgáltam, hogy a szótagszekvenciák hossza összefügg-e a megjelenésük helyzetével a vizsgált beszédegységeken belül. Vizsgáltam az összefüggést a beszédszakaszok szótagszáma és a bennük jelentkező glottalizált szótagok száma között. Elemeztem az irreguláris zöngével megvalósult kitöltött szünetek arányát az összes kitöltött szünethez viszonyítva, ugyancsak beszédhelyzetenként.

A statisztikai elemzést (párosított *t*-próba, Mann–Whitney-próba, Pearson- és Spearman-féle korrelációelemzés) az SPSS 15.0 verziójával végeztem el.

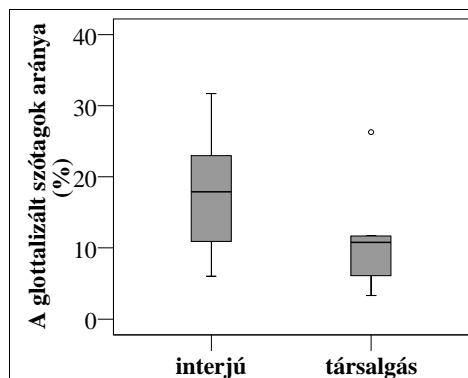
Eredmények

A glottalizáció gyakorisága jelentős különbségeket mutat a két vizsgált beszédhelyzetben (vö. 2. ábra): az interjúban nagyobb arányban jelenik meg irreguláris zöngeminőség, mint a társalgásokban. Az előbbi helyzetben átlagosan a szótagok 17,9%-a volt részben vagy egészben glottalizált, a szórás 9,2%. A társalgásokban az átlag és a szórás is kisebb: 11,5%, illetve 8,0%.

Minden adatközlőre igaz volt, hogy a társalgásban kevesebbet glottalizált, mint az interjúhelyzetben, bár a kettő közötti eltérés mértéke – ahogyan az egyéni glottalizációs gyakoriság is – erősen egyénfüggő. Az adatközlők közül egy 59 éves férfi glottalizált a legkevesebbet, az interjúban 6,0%-os, a társalgásban 3,3%-os gyakorisággal, tehát e két értéke között közel kétszeres szorzó van, bár az értékek igen alacsonyak. A leggyakrabban mindkét helyzetben egy 44 éves nő glottalizált, az interjúban ejtett szótagjainak 31,7%-a volt legalább részben irreguláris, míg a társalgásban ez az arány 26,3% volt, azaz mintegy ötödével kevesebb. A legkisebb eltérést a két beszédhelyzet közötti glottalizáció mértékében egy 60 éves nő beszédében találjuk, aki az interjúban a szótagok 10,9%-át, a társalgásban pedig 10,3%-át ejtette legalább részben irregulárisan.

A két beszédhelyzet közötti glottalizációs gyakoriságbeli eltérések természetesen nem mondanak ellent annak a szakirodalomból is jól ismert ténynek, hogy a beszélőkre jellemző az, hogy milyen mértékben, milyen gyakran glottalizálnak. A Pearson-féle korrelációelemzés eredménye szerint a két be-

szédhelyzetben mért glottalizációs arány között erős, szignifikáns összefüggés van ($r = 0,818$, $p = 0,047$), vagyis aki átlagosan gyakrabban alkalmaz irreguláris zöngét az interjúban, az így tesz a társalgásban is (és fordítva). Az ugyanakkor nem volt kimutatható, hogy a glottalizációs gyakoriság és a két helyzet közötti különbség mértéke között is lenne összefüggés, azaz nem nagyobb mértékű a két helyzet közötti különbség a glottalizációt tekintve, ha a beszélő hajlamosabb a glottalizációra (vagy fordítva).



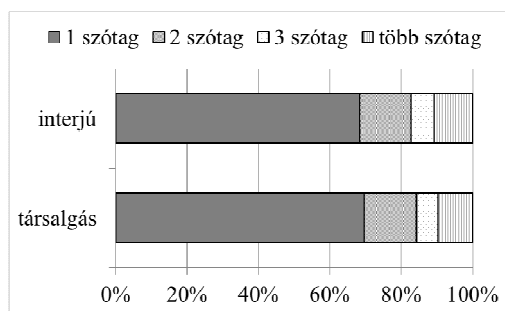
2. ábra

A glottalizált szótágok aránya a vizsgált beszédhelyzetekben

A glottalizált szótagszekvenciák hosszában nem fedezhető fel különbség a kétféle beszédhelyzetben rögzített beszédprodukciók között. Mindkét alkorpuszban az egy szótagnyi irregularitás a leggyakoribb, az előfordulások mintegy 70%-át teszik ki. Jóval kisebb, de ugyancsak egyenlő arányban jelentkeznek két szótagos glottalizált szekvenciák, mintegy 15%-ban. A három és több szótagra kiterjedő irregularitás 16-17%-ot tett ki (3. ábra).

A glottalizáció pozíciójának elemzése előtt érdemes megvizsgálni, hogy a különböző, vizsgált szónál nagyobb, beszédegységek határai hogyan viszonyulnak egymáshoz. Az összes vizsgált határ 16,4%-ában egybeesett a szakasz, a frázis és az elemi mondat eleje, mindhárom beszédegység vége pedig 15,7%-ban vágott egybe. A frázis és az elemi mondat eleje az összes határ 14,6%-ában egyezett meg, és ugyanilyen arányban esett egybe a frázisvég az elemi mondat végével. A beszédszakasz eleje a frázis elejével 4,0%-ban, a szakasz vége a frázis végével 4,4%-ban egyezett meg. A szakaszok eleje nem vágott egybe más beszédegység elejével 10,9%-ban, a szakaszvégek esetében ez az arány 11,3%. Az önálló (más egység határával nem egyező) fráziskezdet és frázisvég egyaránt 3,3%-ot tett ki, és az elemi mondat eleje és vége 0,7-0,7%-ban nem esett egybe más egység határával. Mindebből az a következtetés vonható le, hogy hasonló arányban fordul elő az, hogy a frázis

és/vagy az elemi mondat határa a beszédszakasz belsejébe esik, mint az, hogy a szakaszhatár széttagolja a frázist és/vagy az elemi mondatot. Mindeközben viszonylag ritka, hogy a frázis és/vagy az elemi mondat határai egybevágjanak a beszédszakasz határaival.



3. ábra

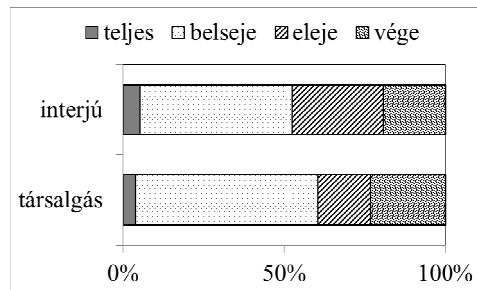
Az irreguláris zöngével realizált szótagok szekvenciáinak hossza a vizsgált beszédhelyzetekben

A beszédshituációk között jelentősebb eltéréseket találunk a glottalizáció beszédszakaszbeli helyzetét illetően (4. ábra). A glottalizáció a teljes szakaszban jelentkezett a szekvenciák 5,4%-ában az interjúkban, 3,9%-ukban a társalgásokban. A glottalizáció a szakasz belsejében jelent meg az interjúbeli szekvenciák 47,1%-ában, a társalgásbeliek 56,6%-ában. A szakasz eleje volt glottalizált az interjúk szekvenciáinak 28,2%-ában, ez az arány a társalgásokban csak 16,2%. A szakasz végi előfordulásokat tekintve ugyanakkor kisebb a különbség: ebben a pozícióban a két beszédprodukcióban átlagosan 19,3%-os és 23,2%-os gyakoriság adódott. Ha együtt vizsgáljuk a határhelyzetet elfoglaló (a szakasz elején és végén megjelenő) glottalizációt, szembeállítva a nem vagy nem csak határon jelentkező irregularitással, az derül ki, hogy az interjúkban az előfordulásoknak közel a fele (47,5%) határpozíciójú volt, a társalgásokban ez az arány kisebb (39,5%).

Elemeztem, hogy van-e összefüggés a beszédszakasz hossza (szótagszáma) és a benne irreguláris zöngével realizálódott szótagok száma között. Mind az interjúkban, mind a társalgásokban szignifikáns, gyenge-közepes összefüggést mutatott ki a Spearman-féle korrelációelemzés ($\rho = 0,334$ és $0,376$, $p < 0,001$). A két helyzetben a beszédszakaszok szótagszáma nem tér el a Mann-Whitney-próba szerint.

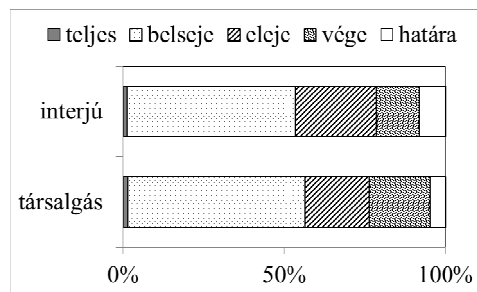
Az elemi mondatokon belüli pozíciókat tekintve az interjú glottalizált szótagsekvenciáinak 52,3%-a, a társalgás glottalizált szótagjainak 55,3%-a elemi mondat belsejében jelent meg. Az interjúkban 25,2%, a társalgásokban 19,7% mondat eleji pozícióban realizálódott, az előbbi beszédhelyzetben

13,2%, az utóbbiban pedig 18,9% mondat végén. Mondathatáron az interjúkban az esetek 8,1%-ában, a társalgásokban 4,8%-ban adatoltam glottalizációt (beleértve ebbe azt is, amikor a glottalizált szótag az elemi mondatokat összekapcsoló kötőszóban jelent meg). Mindkét beszédhelyzetben – várhatóan – azok az esetek voltak a legritkébbak, amikor a glottalizáció az elemi mondat egészében észlelhető volt: 1,2%, illetve 1,3% (5. ábra). Összességében az elemi mondatok szélén (elején, végén és határán) jelentkező glottalizáció a mondat belsejével hasonló arányban volt adatolható mindkét szövegben: az interjúkban 46,5%, a társalgásokban 43,4%.



4. ábra

A glottalizált szótagszekvenciák helyzete a beszédszakaszokon belül

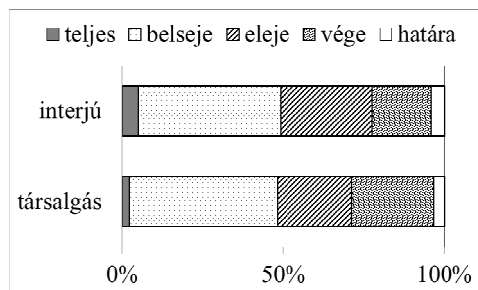


5. ábra

A glottalizált szótagszekvenciák helyzete az elemi mondatokon belül

A spontán beszéd egységeinek hangzó megvalósítása az intonációs frázisok vizsgálatát igencsak megnehezíti, hiszen eltolódnak a hangsúlyok, a frázisokat szünetek szabdalják szét, illetve a viszonylag lapos dallammenetek esetén a frázisok meghatározása olykor ad hoc jelleget ölt. Az intonációs frázisok belsejében jelentkezett a glottalizált szótagok és szótagszekvenciák közel fele az interjúkban (44,0%) és a társalgásokban (46,0%) is. A frázis eleji glottalizáció 28,5%-ot tett ki az előbbi, 23,0%-ot az utóbbi beszédhelyzetben.

A frázis lezárását jelezte az irregularitás az interjúk glottalizált szótagjainak 18,1%-ában, valamint a társalgásbeliek 25,3%-ában. Frázishatárra estek a glottalizált szótagok az interjúban az esetek 4,2%-ában, a társalgásokban 3,4%-ban. A frázis egésze glottalizált volt 5,2%-ban az interjúkban és 2,3%-ban a társalgásokban. Mindkét vizsgált beszédhelyzet esetében hasonló arányban jelentkezett glottalizáció a frázisok határán (50,8% és 51,7%) és a frázisok belsejében (6. ábra).



6. ábra

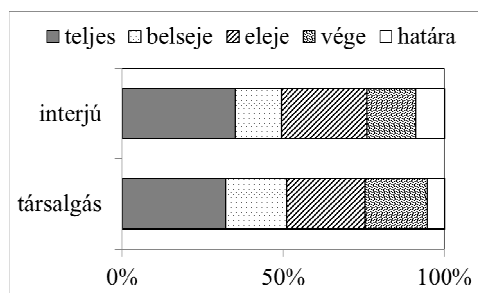
A glottalizált szótagszekvenciák helyzete az intonációs frázisokon belül

A glottalizált szótagok a szavak szintjén nagyjából kiegyenlített pozíciókban jelentek meg mindkét beszédhelyzetben (7. ábra). Az interjúkban a glottalizáció 35,1%-ban érintett teljes szavakat, ez az arány a társalgásokban 32,2%. A szó belseji glottalizáció az előbbi helyzetben 14,3%-ot, az utóbbiban 18,9%-ot tett ki. A szó eleji glottalizáció az interjúkban 26,6%-ban fordult elő, a társalgásokban 24,2%-ban. Az interjúkban a szó vége ritkábban volt irregularis, itt az arány 15,1%, a társalgásokban 19,4%. Szóhatáron átfelölő irregularitást az esetek 8,9%-ban találtam az interjúkban, 5,3%-ban a társalgásokban. Összességében mindkét beszédhelyzetben kiegyenlítettnek mondható a szóhatáron és az egyebütt megjelenő glottalizáció aránya.

Felmerül a kérdés, hogy a különböző pozíciókban milyen hosszúságú szótagszekvenciák jelennek meg – gyakrabban fordulnak-e elő hosszabb szekvenciák például beszédszakasz vagy frázis végén, mint máshol. Az interjúkban (8. ábra) a glottalizált szekvenciák hosszának növekedésével nem tolódik el (jelentősen) az eloszlás a frázis, a beszédszakasz vagy az elemi mondat vége felé. Ugyanakkor a glottalizáció nyúlásával tendenciaszerűen csökken az előfordulások száma mind a négy vizsgált beszédsegység elején, azaz a szavak, szakaszok, frázisok és elemi mondatok elején megjelenő glottalizáció jellemzően csak az első szótagot érinti.

Az egy és két szótagú szavak esetében nagyjából azonos arányban (az esetek mintegy ötödében) glottalizált a teljes szóalak (ezek természetesen főleg kötőszavak, névelők, egyéb funkciószavak). Jelentős ugyanakkor a különbség

abban a tekintetben, hogy a két szótagnyi glottalizáció a legnagyobb arányban (33,3%) szóhatáron átívelve jelentkezik. A három és több szótagnyi glottalizáció esetében – várhatóan – a teljes szó érintett a legtöbb esetben.



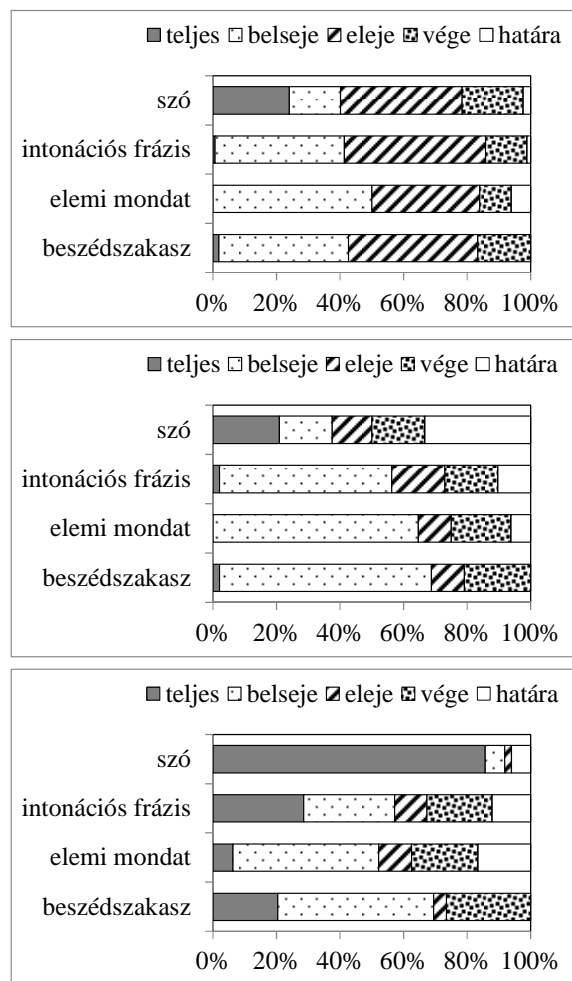
7. ábra

A glottalizált szótagszekvenciák helyzete a szavakban

Míg az egy szótagnyi glottalizáció a szónál nagyobb szinteken nagyjából kiegyenlítetten jelentkezik az intonációs frázis, az elemi mondat és a beszédszakasz belsejében és elején, addig a két szótagnyi glottalizáció túlnyomórészt ezek belsejében fordul elő. A három és több szótagnyi glottalizáció esetében sokkal változatosabb mintázatot kapunk: a frázis szintjén szinte kiegyenlítettek a különböző pozíciók arányai, az elemi mondatban és a beszédszakaszban itt is a belső helyzetű glottalizáció a leggyakoribb. Valamelyest gyakrabban fordul elő ilyen hosszú glottalizáció a beszédszakaszok végén, és a két szótagnyihoz képest az elemi mondat határán is.

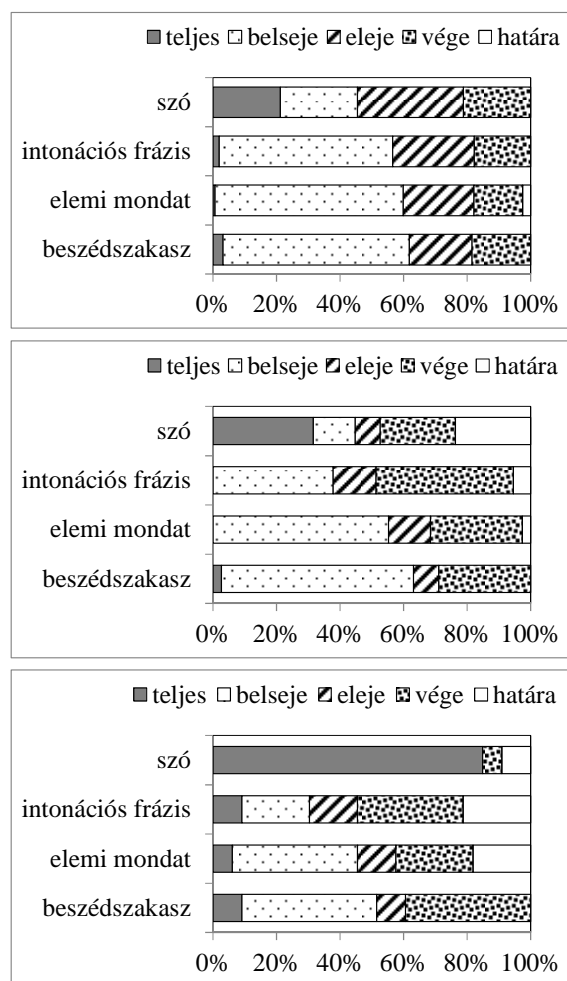
A társalgásokban (9. ábra) szószinten az interjúkban tapasztalhatóhoz hasonló mintázatokat látunk. A szónál nagyobb egységekben az egy szótagnyi glottalizáció 50%-nál nagyobb arányban jelentkezik a frázisok, a beszédszakaszok és az elemi mondatok belsejében, és nagyjából kiegyenlített arányban tapasztalható egy szótagnyi glottalizáció ezen egységek elején és végén. A két szótagnyi glottalizáció elsősorban az intonációs frázisok végére jellemző, kisebb mértékben a belsejére.

A társalgásokban a három és több szótagnyi glottalizáció esetében a frázisbeli pozíciók aránya kiegyenlítettebb, a frázis végi helyzet után a frázis belseje és a frázis határán jelentkező hosszú glottalizáció a leggyakoribb. A beszédszakaszok és az elemi mondatok esetében a két és a több szótagnyi glottalizáció az interjúkhoz hasonlóan alakult: főleg ezek belsejében jelentkezik a glottalizáció, hasonló, illetve valamivel kisebb arányban a végükön. A hosszabb glottalizáció gyakrabban ível át a frázis és az elemi mondat határán.



8. ábra

Az egy szótagot érintő (fent) a két szótagos (középen) és a három és több szótagos glottalizáció (lent) elhelyezkedése a különböző vizsgált beszédegységekben az interjúkban



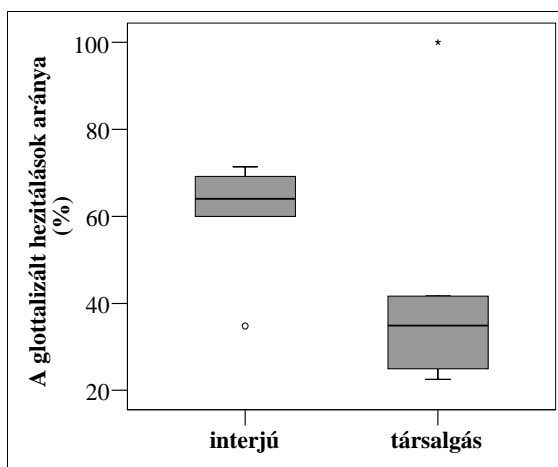
9. ábra

Az egy szótagot érintő (fent) a két szótagos (középen) és három és több szótagos glottalizáció (lent) elhelyezkedése a különböző vizsgált beszédegységekben a társalgásokban

Végül lássuk a kitöltött szünetek eredményeit. Ezek gyakorisága különbséget mutatott a két beszédhelyzetben (ahogy az a szakirodalom alapján várható volt, vö. Bóna 2013). Az interjúkban percenként átlagosan 5,4 kitöltött szünet jelent meg (a szórás 3,3), a társalgásokban 4,6 volt a percenkénti szá-

muk (a szórás 3,0). A kitöltött szünetek viszonylag gyakran realizálódtak részben vagy egészben irreguláris zöngével. A 10. ábra mutatja a glottalizált kitöltött szünetek arányát a produkcióban megjelenő összes kitöltött szünethez viszonyítva a két beszédhelyzetben.

Látható, hogy volt olyan adatközlő, akinek a társalgásbeli kitöltött szünetei 100%-ban glottalizáltak voltak, hozzá kell tenni azonban, hogy ez a beszélő mindösszesen két alkalommal hezitált a társalgásban. Ezért az átlagszámításakor ezt az adatot a továbbiakban nem veszem figyelembe. Így átlagosan a társalgásbeli kitöltött szünetek 31,8%-a volt részben vagy egészben irreguláris, a szórás itt 7,7%. Az interjúban a hezitálásoknak átlagosan 60,6%-a volt irreguláris, itt 8,8% a szórás. A viszonylag látványos különbség ellenére a beszélőnkénti összehasonlításban a párosított *t*-próba nem mutatott ki szignifikáns eltérést a beszédhelyzetek között a glottalizált hezitálások arányát illetően. A korrelációelemzés szerint nincs összefüggés az egyes beszélők hezitálásainak glottalizációs arányát tekintve a két beszédhelyzetben, azaz azt a fajta egyéni mintázatot, amelyet a szövegszótagok esetében láttunk, a kitöltött szünetekre vonatkozóan nem lehet kimutatni.



10. ábra

A glottalizált hezitálások aránya az összes hezitáláshoz viszonyítva beszédhelyzetenként

Következtetések

Kutatásomban a BEA 10 beszélőjének anyaga alapján azt vizsgáltam, hogy két spontán beszédhelyzetben, az interjúban és a társalgásban eltér-e a glottalizáció gyakorisága, valamint pozíciói, illetve ebből adódóan a funkciói.

A két vizsgált beszédhelyzet különbözik abban, hogy a társalgásban a beszédpartnerek fordulói alatt részben lezajlanak a tervezési folyamatok, amelyeket a percepció működése is nagyban elősegítenek, hiszen az adott témában releváns nyelvi egységek aktiválódnak a feldolgozás révén, ezért a megszólaláshoz készülven előhívásuk sokkal könnyebb. Ez a helyzetek közötti eltérés több beszéd-sajátosság különbségeit magyarázza, például a szünetezését (a szünetek helye, időtartama, beleértve a kitöltött szüneteket). Azt feltételeztem, hogy az irreguláris zöngeminőség ritkább a társalgásban, ugyanakkor gyakrabban tölt be határjelző szerepet, mint az interjúhelyzetben.

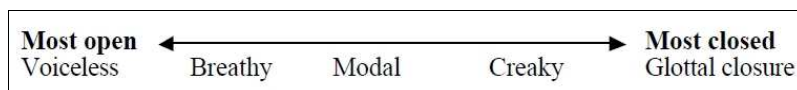
Hipotézisem csak részben igazolódott: a glottalizáció valóban szignifikánsan ritkább volt a társalgásban. Mindemellett az interjúban nem hosszabbak az irreguláris szótagszekvenciák, mint a társalgásban, hiszen a relatíve legtöbb glottalizáció itt is egy szótagnyi egységekben jelentkezett. A különböző hosszúságú szótagszekvenciák relatív gyakorisága ugyancsak azonos volt.

A beszédszakaszok a szövegformálás olyan egységei, amelyeknek a határai a spontán beszédben nem feltétlenül esnek egybe nyelvi egységek (pl. elemi mondatok vagy akár szintagmák) határaival, hiszen a beszédtervezés és -kivitelezés párhuzamos folyamataiból adódóan a megvalósuló megformálás gyakran nem illeszkedik tökéletesen a nyelvi konstrukciókhoz. A beszédszakaszok határait éppen ezért nagyobb mértékben befolyásolhatják olyan jelenségek, mint a levegővétel vagy a tervezési bizonytalanságokból, illetve az önellenőrzési folyamat működtetéséből adódó szünetek. Ebből adódóan a zöngés hangok esetében a beszédszakasz elején történik meg a zöngé felépülése, a szakasz végén pedig a lecsengése. Ladefogednek (1971, idézi Gordon–Ladefoged 2001) a 11. ábrán látható modellje a glottalizációt átmenetként képzelel el a hangszalagoknak a zöngét közvetlenül megelőző és követő zárt pozíciója és a modális zöngére alkalmas fonációs helyzete között. A beszédszakaszok elején és végén (határán) jelentkező glottalizáció tehát ez alapján értelmezhető a zöngé felépülésével és lecsengésével asszociálódó, túlnyomórészt fiziológiai okokkal magyarázható jelenségnek. Az interjúkban az előfordulásoknak közel a fele (47,5%) határpozíciójú volt, a társalgásokban ez az arány csak 39,5%. A viszonylag nagy különbséget az okozza, hogy az interjúkban majdnem kétszer gyakrabban indul irreguláris zöngével a szakasz (28,2%), mint a társalgásokban (16,2%).

Az az eredmény, hogy a beszédszakasz szótagszáma és a benne irreguláris zöngével realizálódott szótagok száma közötti összefüggés gyenge, arra utal, hogy a glottalizáció megjelenését önmagában nem indukálja a szakasz hossza és ezáltal a modális zöngé fenntartásának esetlegesen nehezített volta, pusztán arról van szó, hogy a nagyobb szótagszám esetén potenciálisan több szótag realizálódhat irreguláris zöngével.

Azt a feltételezést, hogy az irreguláris zöngeminőség gyakrabban töltene be határjelző szerepet a társalgásban, mint az interjúhelyzetben, többféle beszédjegységen vizsgáltam. Az elemi mondatok, az intonációs frázis és a sza-

vak esetében a határhelyzetben jelentkező glottalizáció ugyanolyan gyakori volt az interjúkban, mint a társalgásokban, és mindkét esetben kiegyenlített volt a határhelyzeti és a beszédegység belseji előfordulások aránya. A második hipotézis tehát nem igazolódott.



11. ábra

A fonációtípusok kontinuum (Gordon–Ladefoged 2001: 384)

(A két végpont a legnyíltabb és a legzártabb állás; a legnyíltabbhoz tartozik a zöngétlen hangok képzése, innen fokozatosan záródnak a hangszalagok, ami levegős zöngéképzéshez vezet. A hangszalagok mintegy középső állása a modális zöngét idézi elő, ebből a zárt állás felé elmozdulva jelenik meg a glottalizáció. Végül a teljes glottális zárhoz jutunk.)

Azt, hogy a társalgásban a kitöltött szünetek ritkábbak, mint az interjúkban, a szakirodalom (pl. Bóna 2013) azzal indokolja, hogy a tervezés a társalgási helyzetben bizonyos értelemben könnyített az interjúszituációhoz viszonyítva. A hezitálások irregularitását vizsgálva az az eredmény adódott, hogy a társalgásban nemcsak ritkább a hezitálás, de kisebb arányban glottalizált is.

Az eredmények azt igazolják, hogy a fiziológiai sajátosságokból (elsősorban a zöngé felépülése a beszédszakaszokban, másrészt a nem szándékosan artikulált kitöltött szünetek esetében) adódó glottalizáció viszonylag gyakoribb az interjúkban, mint a társalgási helyzetben. A szándékosnak tekinthető határjelzés mintázatai ugyanakkor nem térnek el a beszédhelyzet függvényében. Az interjúkban adódott gyakoribb glottalizáció így összefügghet azzal, hogy ebben a szituációban a beszélő feladata nehezebb, hiszen (az interjúkésztő egy-egy kérdését leszámítva) mindvégig egyidejűleg zajlanak a beszédtervezési és -kivitelezési folyamatok. Lehetséges, hogy a gyakoribb glottalizáció, ezen belül különösen a kemény zöngéindítás és a részben vagy egészben irreguláris kitöltött szünet a tervezési nehézségek öntudatlan lecsapódása az automatikusan működtetett artikulációban. Ugyanakkor szerepe lehet ebben annak is, hogy a formalitási skálán az interjú a formálisabb, a társalgás az informálisabb végponthoz közelebb helyezkedik el, és a glottalizáció a feszebb artikulációs működések eredménye, amelyeket a formális helyzet vált ki. Természetesen sok más ok is számba jöhet, amelyeket komplexebb vizsgálatok tárhatnak fel.

Irodalom

- Batliner, Anton – Burger, Susi – John, Birgit – Kiessling, Andreas 1993. MÜSLI: A classification scheme for laryngealizations. In *Working Papers, Prosody Workshop*. Lund, Sweden. 176–179.
- Boersma, Paul – Weenink, David 2013. *Praat: doing phonetics by computer [Computer program]. Version 5.3.* <http://www.praat.org>. (Letöltés ideje: 2013. júl. 9.)
- Bóna Judit 2013. A beszédsszünetek fonetikai sajátosságai a beszéd típus függvényében. *Beszéd kutatás* 2013. 60–75.
- Bóhm Tamás – Ujváry István 2008. Az irreguláris fonáció mint egyéni hangjellemző a magyar beszédben. *Beszéd kutatás* 2008. 108–120.
- Dilley, Laura – Shattuck-Hufnagel, Stefanie – Ostendorf, Mari 1996. Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics* 24. 423–444.
- E[lekfi] L[ászló] 1983. Hanglejtés. In Grétsy László – Kovalovszky Miklós (főszerk.): *Nyelvművelő kézikönyv. Első kötet. A–K. Második, változatlan kiadás.* Akadémiai Kiadó, Budapest. 755–784.
- Elekfi László 1992. *A magyar hangkapcsolódások fonetikai és fonológiai szabályai.* Linguistica. Series A. Studia et dissertationes 10. MTA Nyelvtudományi Intézet, Budapest.
- Gerratt, Bruce R. – Kreiman, Jody 2001. Toward a taxonomy of nonmodal phonation. *Journal of Phonetics* 29. 365–381.
- Gordon, Matthew Ladefoged, Peter 2001. Phonation types: a cross-linguistic overview. *Journal of Phonetics* 29. 383–406.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya.* Osiris Kiadó, Budapest.
- Gósy Mária – Gyarmathy Dorottya – Horváth Viktória 2009. A beszéd természetességéről alkalmazott fonetikai szempontból. *Beszéd kutatás* 2009. 170–181.
- Gósy Mária – Gyarmathy Dorottya – Horváth Viktória – Grácz Tekla Etelka – Beke András – Neuberger Tilda – Nikléczy Péter 2012. BEA: Beszélt nyelvi adatbázis. In Gósy Mária (szerk.): *Beszéd, adatbázis, kutatások.* Akadémiai Kiadó, Budapest. 9–24.
- Houde, Robert A. – Hillenbrand, James M. 1994 The role of voice pitch in the perception of glottal stops. *Journal of the Acoustical Society of America* 95. (S1) 2872.
- Huber, Dieter 1988. *Aspects of the communicative function of voice in text intonation.* PhD thesis. Chalmers University, Göteborg–Lund.
- Ladefoged, Peter 1971. *Preliminaries to linguistic phonetics.* University of Chicago, Chicago.
- Laver, John 1980. *The phonetic description of voice quality.* Cambridge University Press, Cambridge.
- Markó Alexandra 2005. *A spontán beszéd néhány szuprasegmentális jellegzetessége. Monologikus és dialogikus szövegek összevetése, valamint a hümmögés vizsgálata.* PhD-értekezés. ELTE, Budapest.
- Markó Alexandra 2009. Stigmatizált hanglejtésforma a spontán beszédben. *Beszéd kutatás* 2009. 88–106.
- Markó Alexandra 2010. A prozódia szerepe a spontán beszéd tagolásában. *Beszéd kutatás* 2010. 82–99.
- Markó Alexandra 2011. A glottalizáció határjelző szerepe a felolvasásban. *Beszéd kutatás* 2011. 31–45.

- Markó Alexandra 2012a. Az irreguláris zöngé szerepe a magánhangzók határának jelölésében V(#)V kapcsolatokban. *Beszéd kutatás 2012*. 5–29.
- Markó Alexandra 2012b. Az irreguláris zöngé funkciói és gyakorisága olvasott és spontán beszédben. In Gósy Mária (szerk.): *Beszéd, adatbázis, kutatások*. Akadémiai Kiadó, Budapest. 25–42.
- Markó Alexandra 2013. A magánhangzók határának jelzése glottalizációval VV kapcsolatokban – hivatásos és tipikus beszélők összevetése. In Gecksó Tamás – Sárdi Csilla (szerk.): *Nyelvhasználat a médiában*. Segédkönyvek a nyelvészet tanulmányozásához 155. Kodolányi János Főiskola–Tinta Könyvkiadó, Székesfehérvár–Budapest. 122–127.
- Markó Alexandra – Bóna Judit 2012. Elterő beszéd módok intonációs sajátosságai fiatal és idős korban. In Balázs Géza – Veszelszki Ágnes (szerk.): *Nyelv és kultúra: Kulturális nyelvészet*. Magyar Szemiotikai Társaság, Budapest. 253–258.
- Redi, Laura – Shattuck-Hufnagel, Stefanie 2001. Variation in the realization of glottalization in normal speakers. *Journal of Phonetics* 29. 407–429.
- Tolcsvai Nagy Gábor 2011. *Kognitív szemantika*. Konstantin Filozófus Egyetem, Közép-európai Tanulmányok Kara, Nyitra.

A kutatás első eredményei a Beszéd kutatás 2013 konferencián hangzottak el. Köszönöm Mády Katalin, valamint ismeretlen bírálóm kérdéseit és javaslatait, amelyek hozzájárultak az elemzés végső szempontrendszerének kialakításához.

SZÓIDŐTARTAMOK GYERMEKEK ÉS FELNŐTTEK SPONTÁN NARRATÍVÁIBAN

Horváth Viktória

Bevezetés

A spontán beszéd temporális viszonyait különféle tényezők határozzák meg. Ezek részben univerzálisak (pl. fiziológiai), részben nyelvspecifikusak, de meghatározóak a beszélők egyéni sajátosságai (egyéni temperamentum, érzelmek, életkor) és például a beszédhelyzet is.

Az időviszonyok jellemezhetők többek között a beszédsebességgel. Az artikulációs tempó a kiejtés tiszta idejére eső jelek számát jelenti, a beszédtempó tartalmazza a szüneteket és megakadásokat is.

A temporális szerveződés megragadható különböző méretű beszédegységek időtartamában is. Magyar nyelvű spontán narratívákban akusztikai-fonetikai paraméterekkel is jellemzett egységeket sikerült meghatározni (Gósy et al. 2011). Ezek a tematikus egységek a beszélők többségénél hasonló szerkezettel valósultak meg, az alaphangmagasság és az intenzitás magasabb volt a tematikus egység elején, mint a végén. A tematikus egységek rövidültek a narratíva vége felé. Ez a tény az ún. belső temporális szervező működését valószínűsíti a beszélőknél; amelyet alátámasztott az is, hogy az artikulációs tempó változásaiban bizonyos beszélőfüggetlen mintázatot lehetett kimutatni.

A nagyobb egységek időtartamviszonyai mellett a szavak időtartama is fontos szervezője a beszéd temporális jellemzőinek. A nemzetközi kutatások szerint a szavak időtartama nemcsak a beszéd temporális viszonyairól nyújt információt, hanem a beszédtervezési folyamatok időzítési sajátosságairól is. A szóidőtartamok variabilitását számos tényező befolyásolja, mint például maga a nyelv típusa, a beszédmód, a beszélő személye stb. (Baker–Bradlow 2009; Baker et al. 2011). Maga a szó típusa és a gyakorisága is hatással van az időtartamra: a gyakori tartalmas szavak rövidebbek, mint a hasonló szerkezetű ritkábbak, de a funkciószavak esetében nem igazoltak ilyen jellegű különbséget (Bell et al. 2009). A gyakoribb szavak redukálódását már a 12. században kimutatta egy arab „nyelvész” (Leslau 1969). A gyakoribb szavak nemcsak rövidebb időtartamban valósulnak meg, hanem gyakran redukált magánhangzókat tartalmaznak, és egész szótagok is törölődhetnek (Zipf 1929; Fidelholz 1975; Aylett–Turk 2004; Pluymaekers et al. 2005).

A magyar spontán beszéd szóidőtartamaival csupán néhány kutatás foglalkozott. Ezek azt igazolták, hogy a szavak objektív időtartamát a beszédben

számos tényező befolyásolja. Minél több beszédhangból, illetve szótagból áll egy szó, annál több időre van szükség a kiejtéséhez. A szótagszám növekedésével az időtartam nem lineárisan növekszik (Gósy 1997). Az artikulációs tervezés és kivitelezés bonyolultsága is hatással van a szóidőtartamokra. A fonológiai koartikulációs szabályt tartalmazó szavak hosszabbak, mint az ilyen folyamatok elvégzését nem igénylő alakok (Gósy 1997). Meghatározó az időtartamra nézve a morfológiai felépítés, a megfelelő toldalékok vagy igekötők kiválasztása, tervezése és kivitelezése időtartam-növekedést eredményezhet (Gósy 1997; Bóna 2011). Fontos tényező továbbá a gyakoriság, a gyakran használt szavakat tendenciaszerűen gyorsabban ejtjük, mint a hasonló struktúrájú ritkébbakat (Gósy 1997). Bizonyos szavakról kimutatták továbbá a funkciófüggő realizációt: töltelékelemként rövidebb időtartammal valósulnak meg, mint kötőszói funkcióban (Gósy–Horváth 2010). A beszélő személye is nagymértékben meghatározza az időtartamokat, és maga a beszédmód is: az egy szótagú szavak például hosszabb időtartamúak a spontán beszédben, mint a felolvasásban, a szótagszám emelkedésével a felolvasott szavak lesznek hosszabbak (Olaszy 2010).

A jelen kutatás célja annak vizsgálata, hogy miként alakulnak a szóidőtartamok felnőttek és kisiskolások spontán beszédében. A gyermekek narratíváinak ilyen szempontú elemzése magyar nyelven elsőként történt meg. Hipotéziseink a következők voltak: a) a beszélő életkora befolyásolja a szavak időzítési sajátosságait, b) a szavak morfológiai struktúrája befolyásolja az objektív időtartamot, c) a „kiegyenlítődési” tendencia (bizonyos számú beszédhang után már nem növekszik lineárisan a szó időtartama) már kimutatható a gyermekek spontán beszédében is.

Anyag és módszer, kísérleti személyek

A szóidőtartamok elemzéséhez két korpuszt használtunk. A gyermekek beszédének vizsgálatához narratívákat rögzítettünk. A protokoll hasonló volt a BEA adatbázisban alkalmazotthoz (Gósy et al. 2012): a kísérletvezető előre megadta a témákat a gyermekeknek, ezt követően csak akkor szólalt meg, amikor segítő, a közlést továbbvivő kérdésre volt szükség. A narratívák témája a gyermekek családja, lakókörnyezete, hobbija, iskolai elfoglaltságai voltak. A teljes korpusz időtartama 41 perc (gyermekenként 3–5 perc). A felvételek a megszokott iskolai környezetben, de csendes körülmények között készültek Sony ICD-SX700 típusú hangfelvevővel; a kísérletvezető minden esetben ugyanaz a személy volt.

A kutatáshoz tíz kilencéves gyermek (5 lány és 5 fiú) beszédét rögzítettük. A gyermekeket véletlenszerűen választottuk ki egy budapesti általános iskola harmadik osztályából. Mindannyian egynyelvűek, tipikus beszédfejlődésűek voltak, és egyiküknek sem volt hallásproblémája és beszédhibája.

A felnőttek szóidőtartamainak vizsgálatához a BEA adatbázisból válogattunk narratívákat 10 felnőtt beszélőtől (5 férfi és 5 nő). Az adatközlők életko-

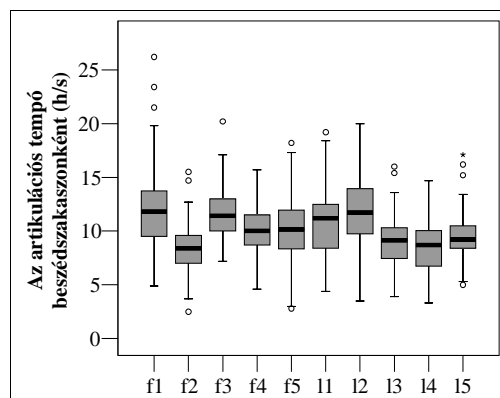
ra 22–35 év. A korpusz időtartama 35 perc (adatközlőként 3–8 perc). Minden beszélő egynyelvű volt, egyiküknek sem volt hallásproblémája és beszédhibája.

A spontán narratívákat beszédszakasz- és szószinten annotáltuk a Praat 5.3 programban (Boersma–Weenink 2013). Automatikusan meghatároztuk a szavak időtartamát és a beszélők artikulációs tempóját. A statisztikai vizsgálatok (regresszióanalízis, Kruskal–Wallis-próba, Mann–Whitney-próba) az SPSS 13.0 programmal történtek.

Eredmények

A gyermekek korpuszában 2215 szót elemeztünk, ezek átlagos időtartama 487 ms (átl. elt.: 387 ms). A toldalékolt szavak aránya 54,8%. A felnőttek korpuszában 3890 szó fordult elő, átlagos időtartamuk 370 ms (átl. elt.: 216 ms). A toldalékolt szavak aránya hasonló, mint a gyermekeknél, 56,4%. (A névelőket nem vettük figyelembe az elemzéskor, hogy az eredmények összevethetők legyenek a magyarra vonatkozó korábbi kutatásokkal).

A beszélő artikulációs tempója befolyásolhatja a szavak időtartamát, ezért beszédszakaszonként elemeztük a tempóértékeket. A gyermekeknél adatolt átlagos artikulációs tempó 10,3 hang/s. Az egyéni különbségek meglehetősen nagyok. A leglassabb átlagos artikulációs tempó 8,5 hang/s, a leggyorsabb pedig 11,9 hang/s, mindkettőt fiúknál adatoltuk (1. ábra).

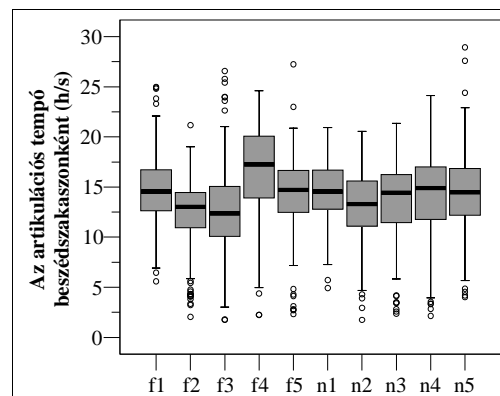


1. ábra

Az artikulációs tempó értékei a gyermekeknél

A felnőttek átlagos artikulációs tempója 13,1 hang/s. A leglassabb átlagos tempó 11,6 hang/s, a leggyorsabb pedig 15,4 hang/s volt (2. ábra). A felnőttek átlagosan három hanggal többet ejtenek másodpercenként a kisiskolásokhoz képest. A különbség statisztikailag is szignifikáns (az adatok nem normál

eloszlása miatt a nemparametrikus Mann–Whitney-próbát végeztük el: $Z = -24,507$, $p < 0,001$).



2. ábra

Az artikulációs tempó értékei a felnőtteknél

A szavak időtartamát minden beszélő esetében normalizáltuk a saját artikulációs tempójához, de a tendenciák hasonlóan alakultak ahhoz, ahogy a következőkben bemutatjuk, ezért a továbbiakban a tempóértékekkel nem foglalkozunk.

Elemeztük, hogy miként alakul a különböző szótagszámú szavak aránya és időtartama a két korcsoport spontán narratíváiban (1. táblázat). A gyermekeknél legnagyobb arányban két szótagú és egy szótagú szavakat adatoltunk. Ezekhez képest a három szótagból álló lexémák fele olyan ritkán fordultak elő (*Erdélybe, fociztunk*). A négy vagy annál több szótagból álló szavak aránya 10% alatti, például: *játékokkal, meglepetésből, számítógépezni, keresztanyukájához*. A maximális szóhossz a kisiskolásoknál hét szótagos volt.

A felnőtteknél az egy szótagos szavak voltak a leggyakoribbak, a két szótagból állók valamivel ritkábbak. A három szótagosok aránya itt is nagyjából a fele az előzőeknek, 10% körüli a négy szótagúak előfordulása. A hosszabb szavak aránya a felnőtteknél is 5% alatti; beszédükben még nyolc (0,2%) és kilenc szótagos (0,1%) lexémákat is adatoltunk, például: *determinisztikusak, számítógép-kezelői, médiainformatikai*.

A különböző szótagszámból álló szavak aránya a két korpuszban nagyon hasonlóan alakult. A hosszabb szavak előfordulásában van némi különbség a felnőttek javára, de a kis elemszámok miatt ez nem jelentős.

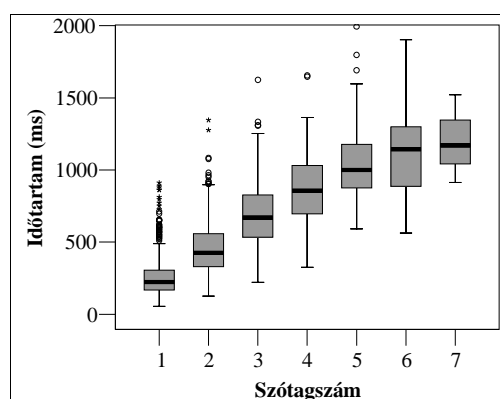
Az időtartamokat elemezve az eredmények azt mutatták, hogy a szótagszám emelkedésével nőtt a szavak átlagos időtartama a gyermekek beszédében (1. táblázat). Az egy szótagból álló lexémák átlagos időtartama 258 ms, az öt, vagy annál szótagból állók kiejtéséhez pedig átlagosan egy másodperc-

nél is hosszabb idő szükséges. Ugyanez a tendencia volt kimutatható a felnőtteknél is: az egy szótagos szavak átlagos időtartamához képest a hét szótagosok átlaga mintegy ötször volt hosszabb. A nyolc szótagosok átlagos időtartama 1033 ms, a kilenc szótagosoké pedig 1165 ms.

1. táblázat: A különböző szótagszámú szavak aránya és időtartama a gyermekek és felnőttek korpuszában

Szótag-szám	Arány (%)		Átlagos időtar-tam (ms)		Átlagos eltérés (ms)	
	Gyerek	Felnőtt	Gyerek	Felnőtt	Gyerek	Felnőtt
1	34,2	35,0	258	200	130	94
2	35,8	31,7	454	328	171	112
3	18,0	17,1	691	488	213	133
4	8,7	9,9	872	627	234	151
5	2,5	4,2	1049	754	287	182
6	0,7	1,4	1134	853	364	195
7	0,1	0,4	1202	1024	305	240

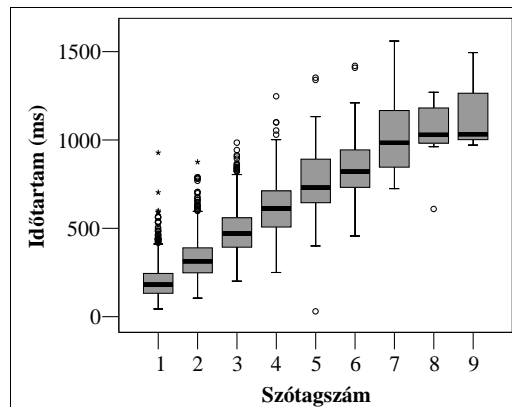
A statisztikai elemzés azt mutatta, hogy a szótagszám emelkedésével az időtartamok szignifikánsan hosszabbak a gyermekek beszédében (Kruskal–Wallis-próba: $\chi^2 = 1370,703$; $p < 0,001$). Ez a szignifikáns különbség azonban csak az öt szótagos szavakig kimutatható a csoportok között (minden esetben $p < 0,001$), az öt, hat és hét szótagos lexémák időtartama már csak tendenciaszerűen különbözik egymástól (3. ábra).



3. ábra

A különböző szótagszámú szavak időtartamai a gyermekeknél

A felnőttek beszédében is igazolható (4. ábra), hogy a szavak időtartama szignifikánsan különbözik egymástól a szótagszám függvényében ($\chi^2 = 2580,625$; $p < 0,001$). A csoportok közötti eltérés (minden esetben $p < 0,001$) a hét szótagos szavakig szignifikáns mértékű, ezt követően már csak tendenciaszerű.



4. ábra

A különböző szótagszámú szavak időtartamai a felnőtteknél

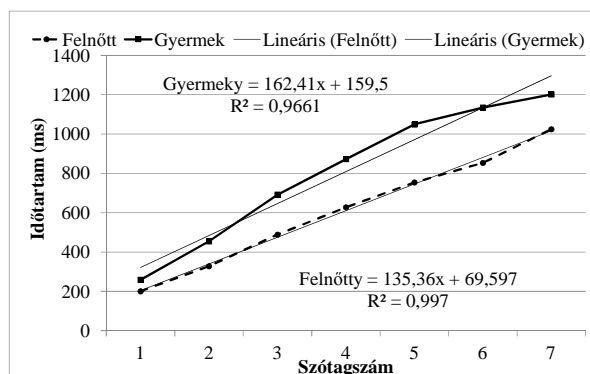
A két korcsoport szóidőtartamait összehasonlítva az eredmények azt mutatták, hogy a felnőtteknél mért értékek minden szótagszám esetében kisebbek voltak. A különbség statisztikailag szignifikáns (Mann–Whitney-próba, egy szótagos szavak: $Z = -11,625$ $p < 0,001$; két szótagos: $Z = -17,685$ $p < 0,001$; három szótagos: $Z = -15,942$ $p < 0,001$; négy szótagos: $Z = -12,192$ $p < 0,001$). A statisztikai elemzést az elemszámok miatt csak a négy szótagos szavakig végeztük el, de az ennél hosszabb szavak időtartamai is a felnőtteknél rövidebbek (vö. 1. táblázat). Az 5. ábra a gyermekek és felnőttek szóidőtartamainak átlaga (vö. 1. táblázat) közötti különbségeket szemlélteti a szótagszám függvényében. Az egy szótagból álló lexémák esetében 58 ms az eltérés a két csoportnál adatolt átlagok között. A két szótagos szavaknál a különbség ennek már kétszerese, a három szótagosoknál pedig már majdnem a négyszerese. A szavak szótagszámának emelkedésével folyamatosan nő a két csoportnál mért átlagok különbsége, a legnagyobb eltérés az öt szótagos szavaknál volt kimutatható (295 ms). Az ennél hosszabb lexémáknál csökken az átlagok közötti eltérés.

A regresszióanalízis megerősítette, hogy a gyermekeknél nagyobb mértékű a szavak időtartamának növekedése a szótagszám emelkedésével (náluk a regressziós egyenes meredeksége jóval nagyobb mértékű, vö. 6. ábra).



5. ábra

A két korcsoport különböző szótagszámú szavainak átlagos időtartama közötti eltérés (ms)



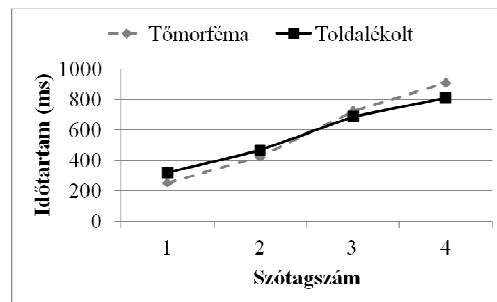
6. ábra

A szótagszám és az időtartam összefüggése a két korcsoportban

A szótagszám mellett a morfológiai struktúra is befolyásolja a szavak időtartamát. A gyermekeknél az egy szótagból álló toldalékolt szavak átlagosan hosszabbak (319 ms, átl. elt.: 141 ms), mint az ugyanannyi szótagból álló tömorfémák (255 ms, átl. elt.: 128 ms); a különbség statisztikailag szignifikáns (Mann-Whitney-próba: $Z = -3,554$ $p < 0,001$, vö. 7. ábra). A két szótagosoknál szintén a morfológiailag strukturált szavak átlagidőtartama hosszabb (468 ms, átl. elt.: 169 ms), mint a tömorfémáké (423 ms, átl. elt.: 170 ms), a különbség szignifikáns ($Z = -3,684$ $p < 0,001$). A hosszabb lexémák esetében a tendencia megfordul, és a több szótagból álló tömorfémák átlagosan hosszabbak, mint a toldalékolt alakok, de a különbség nem szignifikáns. A három szótagos tömorfémák át-

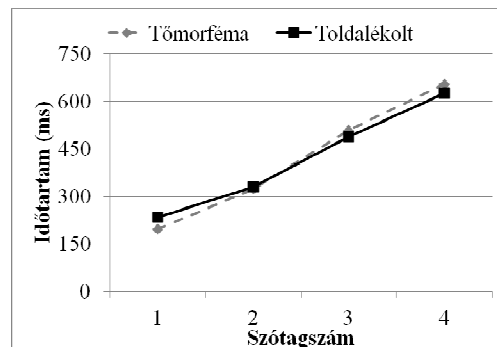
lagos időtartama 724 ms, a toldalékoltaké 689 ms (átl. elt.: 237 ms és 212 ms). A négy szótagosok esetében a tövek és a morfológiailag strukturált lexémák átlagos időtartama között 100 ms a különbség (910 ms és 810 ms, átl. elt.: 303 ms és 230 ms).

A felnőtteknél az egy szótagos szavaknál ugyanaz a tendencia, mint a gyermekeknél: a toldalékmorfémák alakok hosszabbak (234 ms, átl. elt.: 87 ms), mint a tövmorfémák (197 ms, átl. elt.: 93 ms), a különbség szignifikáns ($Z = -4,317$ $p < 0,001$, vö. 8. ábra). A két szótagos szavak átlagidőtartamában nincs különbség a morfológiai felépítéstől függően (tövmorfémák átlag: 324 ms, toldalékoltak: 329 ms; átl. elt.: 120 ms és 109 ms). A három és négy szótagból álló lexémák esetén ugyanaz a tendencia, mint a gyermekeknél: a tövmorfémák átlagosan valamivel hosszabbak, mint a toldalékolt alakok. A három szótagos tövmorfémák átlagos időtartama 508 ms, a toldalékolt alakoké 487 ms (átl. elt.: 149 ms és 132 ms). A négy szótagos tövek átlagos időtartama 655 ms, a toldalékolt szavaké pedig 625 ms (átl. elt.: 162 ms és 150 ms).



7. ábra

A morfológiai struktúra és az időtartam összefüggése a gyermekeknél



8. ábra

A morfológiai struktúra és az időtartam összefüggése a felnőtteknél (ms)

Következtetések

A jelen kutatásban a szavak időzítési sajátosságait elemeztük a beszélő életkora, a szó szótagszáma és a morfológiai struktúra függvényében felnőtt és – magyar nyelven elsőként – kisiskolások spontán beszédében. Vizsgáltuk továbbá a különböző szótagszámú lexémák arányát is. Egy korábbi kutatás szerint az egy szótagos szavak majdnem 50%-ban vannak jelen a spontán beszédben, a két szótagból álló lexémák aránya 28% (Szende 1973). A jelen kutatás szerint a kilencéves kisiskolások narratíváiban az egy és két szótagból álló szavak aránya 30–35%, a három szótagosoké 20% alatti. A felnőttek és az idősek spontán beszédére vonatkozó korábbi kutatások a gyermekekéhez hasonló értékeket igazoltak (vö. Gósy 1997; Bóna 2011), tehát a különböző szótagszámú szavak aránya a spontán beszédben a beszélő életkorától függetlenül hasonlóan alakul. A szótagszám növekedésével csökken az előfordulási gyakoriság, a hat szótagos lexémáké csupán 1% körüli a gyermekek beszédében.

Hipotézisünknek megfelelően az eredmények igazolták, hogy a beszélő életkora befolyásolja a szóidőtartamokat. A szavak időtartama szignifikánsan hosszabb a kilencéves iskolásoknál, mint a felnőttek beszédében. Ez egyrészt az artikulációs tempó különbségéből adódik: a gyermekek átlagosan három hanggal ejtettek kevesebbet másodpercenként. A szóidőtartamok különbsége az időtartam mellett a beszédtervezési folyamatra vezethető vissza. A mentális lexikonból a szótöveket aktiválni kell, ezekhez a szerkezetnek megfelelő toldalékmorfémát előhívni, a koartikulációs szabályokat érvényesíteni, mindezt az aktuális kivitelezéssel párhuzamosan kell elvégezni (Levelt 1989). Ehhez a többlépcsős művelethez a több beszédtapasztalat miatt a felnőtteknek feltételezhetően kevesebb időre van szükségük. Az időseknél meghatározott átlagértékek közelebb vannak a gyermekek, mint a felnőttek átlagos időtartamaihoz (Bóna 2011). Ez feltehetően azzal magyarázható, hogy az időseknél csökken az artikulációs tempó a fiatal felnőttekhez képest, lassulnak a kognitív folyamatok is, nehezebbé válhat a szóelőhívás (Bóna 2011).

A szavak szótagszámának emelkedésével mindkét csoportban nőtt a szavak átlagos időtartama. Az időtartam-emelkedés szignifikáns mértékű, de a kilencéveseknél csak öt, a felnőtteknél hét szótagig. Ezt követően az időtartamok növekedése már csak tendenciaszerű, megjelenik a kiegyenlítődési tendencia, emiatt bizonyos szótagszám felett nem nő szignifikánsan a szavak időtartama.

A szavak szótagszámának emelkedésével folyamatosan nő a gyermekek és felnőttek szavainak átlagos időtartamai közötti különbség, a legnagyobb eltérés az öt szótagos szavaknál volt kimutatható (295 ms). Az ennél hosszabb lexémáknál csökken az átlagok közötti eltérés. Ez arra vezethető vissza, hogy a regresszióanalízis szerint a gyermekek beszédében a szótagszám emelkedésével nagyobb mértékben nő a szavak időtartama.

Igazoltuk azt a hipotézist, hogy a morfológiai struktúra is befolyásolja a szavak időtartamát. Az egy szótagú toldalékolt szavak szignifikánsan hosz-

szababb, mint az egy szótagú tömorfémák (a beszélő életkorától függetlenül). Ha egy ő, ez, kés stb. szótőhöz még tárgyragot kell rendelni, az növeli az előhívásra és kiejtésre fordított időt. A gyermekek beszédében ez a két szótagos lexémákra is igaz, a felnőtteknél már nincs különbség a két szótagos szavaknál a morfológiai struktúra függvényében. A három, vagy annál több szótagból álló szavak esetén fordított a tendencia: a toldalékolt adatok átlagosan rövidebb időtartamban valósulnak meg, mint a szótövek. Itt már feltételezhetően erősebb a gyakorisági tényező szerepe: a hosszabb szótövek (*Olívia, Dunabogdány*) előfordulása jóval ritkább, mint az azonos szótagszámú toldalékolt alakoké (*elutaztunk, emeletes*). A gyakoribb alakokat tendenciaszerűen rövidebben ejtik a beszélők, mint az azonos szótagszámú ritkábbakat (Gósy 1997). Egy korábbi kutatás fiatal felnőtt beszélőinél ugyanezt a tendenciát igazolták; az idős beszélőknél azonban – a szótagszámtól függetlenül – minden esetben a toldalékolt szavak voltak hosszabb időtartamúak, mint a tömorfémák. Ez arra utal, hogy a kognitív folyamatok lassulása miatt az időseknek több időt kell fordítaniuk a toldalékmorfémák tövekhez rendelésére és artikulációs megvalósítására.

Irodalom

- Aylett, Matthew – Turk, Alice 2004. The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech* 47. 31–56.
- Baker, Rachel E. – Bradlow, Ann R. 2009. Variability in word duration as a function of probability, speech style and prosody. *Language and Speech* 52. 391–413.
- Baker, Rachel E. – Baese-Berk, Melissa – Bonnasse-Gahot, Laurent – Kim, Midam – van Engen, Kristin J. – Bradlow, Ann R. 2011. Word durations in non-native English. *Journal of Phonetics* 39. 1–17.
- Bell, Alan – Brenier, Jason M. – Gregory, Michelle – Girand, Cynthia – Jurafsky, Dan 2011. Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language* 60. 92–111.
- Boersma, Paul – Weenink, David 2013. *Praat: doing phonetics by computer [Computer program]. Version 5.3.* <http://www.praat.org>. (Letöltés ideje: 2013. júl. 9.)
- Bóna Judit 2011. A szavak időzítési sajátosságai idősek spontán beszédében. In Gecső Tamás – Sárdi Csilla (szerk.): *Nyelvi funkciók – stílus és kapcsolat. Segédkönyvek a nyelvészet tanulmányozásához* 139. Tinta Könyvkiadó, Budapest. 39–45.
- Fidelholz, James 1975. Word frequency and vowel reduction in English. *Chicago Linguistic Society* 11. 200–213.
- Gósy Mária 1997. A szavak időzítési sajátosságai a spontán beszédben. *Beszédkutatás* '97. 39–49.
- Gósy, Mária – Horváth, Viktória 2010. Changes in articulation accompanying functional changes in word usage. *Journal of the International Phonetic Association* 40/2. 135–161.
- Gósy Mária – Beke András – Horváth Viktória 2011. Temporális variabilitás a spontán beszédben. *Beszédkutatás* 2011. 5–31.

- Gósy Mária – Gyarmathy Dorottya – Horváth Viktória – Grácz Tekla Etelka – Beke András – Neuberger Tilda – Nikléczy Péter 2012. BEA: Beszélt nyelvi adatbázis. In Gósy Mária (szerk.): *Beszéd, adatbázis, kutatások*. Akadémiai Kiadó, Budapest. 9–24.
- Leslau, Wolf 1969. Frequency as determinant of linguistic change in the Ethiopian languages. *Word* 25. 180–189.
- Levelt, Willem J. M. 1989. *Speaking. From Intention to Articulation*. MIT Press, Cambridge.
- Olaszy Gábor 2010. A beszéd komplex szerkezete. In Németh Géza – Olaszy Gábor (szerk.): *A magyar beszéd. Beszédkutatás, beszédtechnológia, beszédinformációs rendszerek*. Akadémiai Kiadó, Budapest. 9–18.
- Pluymaekers, Mark – Ernestus, Mirjam – Baayen, Harald R. 2005. Lexical frequency and acoustic reduction in spoken Dutch. *Journal of the Acoustical Society of America* 118. 2561–2569.
- Szende Tamás 1973. *Spontán beszédanyag gyakorisági mutatói*. Nyelvtudományi Értekezések 81. Akadémiai Kiadó, Budapest.
- Zipf, George Kingsley 1929. Relative frequency as a determinant of phonetic change. *Harvard Studies in Classical Philology* 15. 1–95.

A tanulmány a Bolyai János Kutatási Ösztöndíj és a 108762 számú OTKA-pályázat támogatásával készült.

GYERMEKEK NEMÉNEK ÉS ÉLETKORÁNAK MEGHATÁROZÁSA A BESZÉDÜK ALAPJÁN

Tóth Andrea

Bevezetés

Az alkalmazott fonetika egy sokat kutatott területe a beszélői profilalkotás, melynek során a beszéd azon akusztikai jellemzőit keresik, amelyek képesek jellemezni a beszélő egy adott tulajdonságát, például életkorát, nemét, pszichikai és érzelmi állapotát vagy testalkatát (Lass–Davis 1976; Amir et al. 2012). Köznapi tapasztalataink is azt mutatják, jóllehet korlátozottan, de képesek vagyunk ezeknek a paramétereknek a becslésére. A beszéd tehát hordoz olyan egyedi akusztikai jellemzőket, amelyek utalnak a produktum létrehozójára (Gósy 2001).

Fiziológiailag az emberi hangképzés alapja a kilégzés okozta levegőáramlás, a hangszalagok fonációs beállítottsága és rezgése, illetve a toldalékcső rezonátortevékenysége (Gósy 2004). A beszédképzésben részt vevő szervek (tüdő, gége, toldalékcső) az életkor előrehaladtával változnak, ami a beszéd akusztikai lenyomatának változását eredményezi (Gósy 2004). Kisgyermekkorban a gége nagysága, a hangmagasság, a hangterjedelem és az átlagos beszédhangképzés tekintetében nincs különbség a nemek között (Balázs 1993).

A beszédképzés során a zöngé a vokális traktusban „formálódik tovább”, amely rezonátorüregként funkcionál. A rezonátorterek által felerősített felhangok a formánsok, amelyek meghatározzák a beszédhangok szerkezetét (Gósy 2004). A formánsértékek az artikulációs gesztusokkal, illetve a toldalékcső alakjával mutatnak összefüggést. Az első formáns az állkapocs nyitáshoz kapcsolódik és a nyelvállás fokával, a második formáns a nyelv vízszintes mozgásával és az ajakműködéssel áll kapcsolatban (Kassai 2002). A vokális traktus növekedésének mértéke és tempója az anatómia és a fonetika sokat kutatott területe (l. például Vorperian et al. 2009). Az anatómiai változás fontos vetülete a beszéd akusztikumában végbemenő változás, amely például a nembeli eltérések kapcsán válik fontossá. Több kutatás tanúsága szerint (Lee et al. 1999; Vorperian et al. 2009) a pubertáskor előtt nincs különbség a vokális traktus hosszában a nemek között. Perry és munkatársai (2001) ezzel szemben azt találták, hogy a beszéd akusztikumában a nembeli eltérések már négyéves korban is kimutathatók, és ezek a testmérettel is kapcsolatban vannak. Az életkor előrehaladtával a formánsértékek csökkennek (Huber et al. 1999; Lee et al. 1999). A formánsok értékei a gyermekeknél a legmagasabbak és a felnőtt férfiaknál a legalacsonyabbak (Huber et al. 1999). A nemek

tekintetében a fiúk formánsértékei alacsonyabbak a velük azonos korú lányokénál (Lee et al. 1999).

Az akusztikai magánhangzótér a gyermekeknél a legnagyobb és a legmagasabb (Deme 2012). A magyar nyelvben a magánhangzókat az első két formánsuk értékével határozzák meg, míg az F_3 értéke a beszélőre jellemző paramétereket hordozhat (Gósy 2004).

Az életkor és a nem beszédalapú meghatározásában a legtöbb fonetikai kutatás az alaphangmagasság és a magánhangzóformánsok szerepét vizsgálja. Hasek és munkatársai (1980) 5 és 10 éves korú gyermekek beszédének vizsgálatakor azt találták, hogy 5–6 éves korban még nincs eltérés az f_0 -ban a nemek szerint, 7 és 10 éves kor között viszont a fiúk alaphangmagassága alacsonyabb volt a velük azonos korú lányokénál. Whiteside és Hodgson (1999) hasonló eredményre jutott 6 és 10 éves korú gyerekek beszédének vizsgálatakor: a lányoknál az alaphangmagasság csökkenése 6 és 10 éves kor között, míg a fiúk esetében 8 és 10 éves kor között jelentkezett. Bennett (1983, idézi Whiteside et al. 2002) longitudinális elemzésében 8 és 11 éves kor között vizsgálta az f_0 változását: nem talált eltérést a nemek között. Perry és munkatársai (2001) 4, 8, 12 és 16 évesek magánhangzóinak vizsgálatakor azt találták, hogy a nemeket jellemző alaphangmagasság-változás csak 12 éves korban jelenik meg a beszédben, ugyanakkor a magánhangzók formánsértékei már gyermekkorban is mutatják a nembeli eltéréseket, az F_3 értékei már 4 éves korban is. Busby és Plant (1995) kutatásának tanúsága szerint a nemek beszédalapú megkülönböztetését 4 és 12 éves kor között elsődlegesen a formánsok, 12 éves kor felett pedig az f_0 biztosítja. A nemek tekintetében az F_2 értékében általánosan, míg az F_1 értékében csak az alsó nyelvállású magánhangzóknál ([a], [æ]) találtak eltérést, ezért feltételezték, hogy ezek a magánhangzók érzékenyebbek a vokális traktus életkor- és nembeli változásaira, mivel képzésükkor a szájüreg nyílt, a szájnyílás pedig tág. Lee és munkatársai (1999) eredményei szerint a formánsfrekvenciák csökkenése a fiúk esetében 11–15 éves kor között, a lányoknál 7–12 éves kor között következik be, ami arra enged következtetni, hogy a vokális traktus növekedése a lányoknál előbb fejeződik be.

Traunmüller és Bezooijen (1994) kutatásának tanúsága szerint 5-ről 11 éves korra 19%-ról duplájára nő a nem felismerésének aránya, és a normál és suttogott ejtésű beszédminták felismerési arányai között csak minimális a különbség. Az azonosításban a szerző szerint a formánsok és az f_0 mellett szupraszegmentumok (például beszédhang) is szerepet játszhatnak. Amir et al. (2012) 8 és 10 éves gyerekek beszédének elemzésekor azt találta, hogy a nemek felismerési aránya 70%-ról 80%-ra nő az életkor előrehaladtával. Az azonosítást segítő akusztikai jegyek közül a szupraglottális térhez köthető jellemzőket tartotta dominánsnak.

Lass és munkatársai (1976) kutatásukban suttogott, normál ejtésű és szűrt magánhangzók segítségével vizsgálták a nemek beszédalapú meghatározását

felnőtteknél. Azt találták, hogy a normál és a szűrt magánhangzók alapján több mint 90%-ban, míg a suttogott ejtésűek alapján 75%-ban különíthetők el a férfi és női beszélők. Ebből arra következtek, hogy az alaphangmagasság elsődleges a vokális traktus rezonáns tevékenységéhez képest. Schwartz (1968) kísérletében a válaszadók a véletlennél jobb arányban voltak képesek a nem megbecslésére zöngétlen frikatívák alapján. Az azonosításban a hangképző rendszer érése mellett szociolingvisztikai, tanult minták is szerepet játszanak (Whiteside–Hodgson 1999).

Az életkor becslésére vonatkozó kutatások változatos eredményeket hoztak: Amir és munkatársai (2012) kísérletükben az életkor becslése 37%-ban volt sikeres 8–18 évesek beszéde alapján. A felnőtt beszélőkkel készített kutatások a következő eredményeket hozták: Gocsál elemzésében (1998) következetes, de pontatlan becslésekről írt. Gósy (2001) kutatásában az életkor meghatározása 70%-ban, Ptacek és Sander (1966) kísérletében 78%-ban volt sikeres. Gorham-Rowan és Laures-Gore (2006) szerint a fiatal és idős beszélők elkülönítése az időskori beszéd sajátosságai miatt lehetséges (például az alaphangmagasság csökkenése, a hangtartás rövidülése, a hangerő csökkenése). A formánsfrekvenciák és az alaphangmagasság mellett vizsgálták még a beszédtempó (Gósy 2001) szerepét is az életkor meghatározásában.

A nem és az életkor becslésére vonatkozó kutatások módszertani, korpuszbeli és résztvevői oldalainak változatossága miatt az eredmények kevésbé összevethetők. A férfi és női beszédhangok eltérései részben magyarázhatók az anatómiai különbségekkel, ugyanakkor a paralingvisztikai jegyek, a tanult genderspecifikus viselkedésminták és azok akusztikai vetületei, például a nőkre jellemzőbb levegős hangképzés (Childers–Wu 1991), az artikuláció pontossága (Diehl et al. 1996) vagy a gége leszorításával elért mélyebb alaphang a fiatal fiúknál (Huszár 2009) sem hagyhatók figyelmen kívül. A testalkat szintén befolyásoló tényező lehet (Xue et al. 2010).

A beszélő nemének és korának meghatározása az automatikus osztályozó rendszerekben is fontos lépés, mivel a nemfüggő modellek működése pontosabb a nemfüggetleneknél (Rao–Nagesh 2011). Az osztályozók felhasználása sokrétű, l. például multimédiás alkalmazásokban (Harb–Chen 2005) és az ember-gép interakcióban (Feld et al. 2010). A beszélő nemének és életkorának meghatározásakor elterjedt az alaphangmagasság (Hu et al. 2012) és a hallás átviteli karakterisztikáját modellező ún. MFC-együtthatók (Melfrekvenciás kepsztrális együtthatók) használata (Rao et al. 2011). A nem és az életkor gépi osztályozásával számos kutatás foglalkozott (l. például Rao et al. 2011; Li et al. 2013). Az alaphangmagasság bevonásával készült modellek a beszélő nemét közel 100%-ban képesek azonosítani (Hu et al. 2012).

Az anatómiai különbségek akusztikai vetületeiből látható, hogy az alaphangmagasság és a formánsfrekvenciák összefüggenek az életkori és a társadalmi nembeli sajátosságokkal a felnőttek esetében. Az óvodás (5 éves) és általános iskolás korú (10 éves) gyermekeknél a vokális traktus hosszában be-

következő változások 12 éves kor előtt nem magyarázzák a nemek akusztikailag lehetséges elkülönítését; Vorperian és társai (2009) kutatása a vokális traktus más (például orális és faringális) részeinek fejlődését tartja befolyásolónak az anatómiai változások közül.

Kutatásomban a magyar nyelvre vonatkozóan elsőként foglalkozom gyermekek nemének és életkorának beszédalapú becslésével. A felnőttkort célzó kutatások alapján feltételeztem, hogy 1. a gyermekek neme és életkora a beszéd alapján meghatározható; 2. a felismerés pontossága javul az életkor előrehaladtával.

Anyag, módszer, kísérleti személyek

A jelen kutatásban 14 óvodás (életkoruk 5 év) és 14 általános iskolás (életkoruk 10 év) spontánbeszéd-anyagát használtam fel, amelyet óvodai és iskolai környezetben rögzítettem, egy Sony ICD-SX800 típusú digitális diktafonnal. Mindannyian ép halló, fővárosi, magyar anyanyelvű gyermekek voltak. A 10 éves korosztály kiválasztásának alapja az volt, hogy a serdülőkori (fiziológiai) változások ebben az életkorban kezdődnek el, és ezek a beszédprodukciónak is hatással vannak (Gósy 2004).

Minden gyermektől ugyanazon képsorozat leírását rögzítettem. Az így rögzített, átlagosan 1 perces hanganyagokból kiválasztottam egy-egy 7 s-os részletet, amely a percepció teszt anyagát képezte. A percepció tesztben 30 egyetemi hallgató (28 nő és 2 férfi) vett részt, átlagéletkoruk 23,5 év volt. A hanganyagot csoportosan, zárt térben, számítógépről lejátszva, aktív hangfalon keresztül hallgatták meg. A hanganyagokat minden esetben ugyanolyan sorrendben játszottam le. A kísérleti személyek feladata az volt, hogy egyszeri meghallgatás után határozzák meg a beszélő nemét és életkorát (óvodás vagy iskolás).

A hanganyagot kézi címkézéssel, hangszinten annotáltam a Praat 5.3 szoftver segítségével (Boersma–Weenink 2013). A percepció tesztben használt 7-7 másodperces hanganyagok beszédhangot tartalmazó részein automatikusan, 0,01 másodpercenként kinyertem a beszélő alaphangmagasságát, amelyet 95%-on újrászűrtem – vagyis a sorba rendezett szám adatok alsó és felső 2,5%-át figyelmen kívül hagytam –, az esetleges zajok kizárása végett. A magánhangzók formánsainak mérését a hanganyagban leggyakrabban előforduló hét magánhangzó esetében végeztem el, ezek az [ɔ], [a:], [ɛ], [e:], [i], [o], [u] voltak. A magánhangzók előfordulási arányai közel azonosak voltak. A formánsok mérésére egy erre a célra írt szkriptet használtam, amely a hangidőtartam 50%-án elhelyezett mérőpontban adta meg az F_1 , F_2 és F_3 értékét. A kinyert adatokat ebben az esetben is újrászűrtem, 95%-on.

A statisztikai elemzést az SPSS 13.0 program segítségével végeztem el. Az elemzett akusztikai jellemzők szerepét egytényezős varianciaanalízis (ANOVA) segítségével elemeztem, minden esetben 95%-os konfidenciaszinten.

A nem és az életkor felismerésének tesztelésére létrehoztam egy gépi osztályozót is. Az osztályozáshoz MFC-együtthatókat és GMM-SVM algoritmust használtam. Az osztályozót a Matlab szoftverben valósítottam meg.

A beszélfelismerésben az MFCC-k kiszámítását a beszéd előfeldolgozása előzi meg, amelynek során előkiemelést hajtanak végre (a jel/zaj viszony javítására), illetve kiszűrik a beszédjelet nem tartalmazó részeket, ezt beszéddektálásnak nevezik (Beke 2012). Az előfeldolgozást a jellemzőkinyerés követi, ami a jelen kutatásban az MFC-együtthatók kiszámítását jelenti. Ennek részeként a beszédjelen gyors Fourier-transzformációt (FFT) hajtanak végre, ezt egy ún. ablakfüggvénnyel súlyozzák, így megkapják 10 ms-onként a beszédjel spektrumát. Második lépésként a kapott spektrumot szűrősoros elemzésnek vetik alá (az ún. Mel-skála alapján), amely a hallás kritikus sávjainak megfelelő érzékenységet követ (Beke–Szaszák 2009). A kapott jellemzők eloszlásából kevert Gauss-modellekkel (GMM) becsültem meg a mintaközepet (átlag) és szórását (Reynolds 1995).

A gépi osztályozásra SVM-et használtam. Az SVM bináris osztályozási problémák megoldásában használt matematikai konstrukció, amely szupportvektorok segítségével ún. döntési hipersíkot határoz meg, és így különíti el a két osztály mintáit (Beke 2011). Az osztályozás két részből áll: a tanítás során modelleket készítenek – a jelen kutatásban fiú-lány, illetve óvodás-iskolás modelleket –, majd a teszteléskor az osztályozó a bejövő beszédmintát összehasonlítja a tanító modellel, azaz mintaillesztést hajt végre.

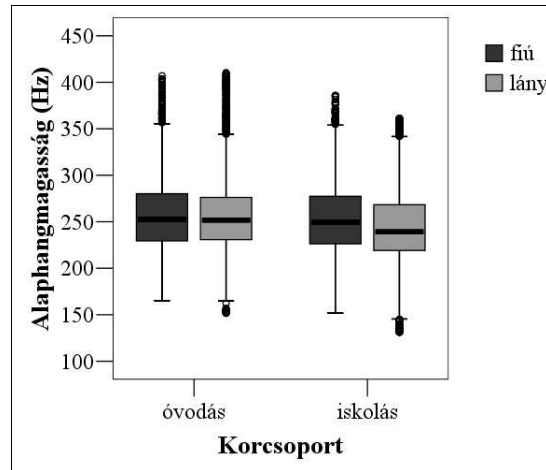
Eredmények

Az akusztikai elemzés eredményei

A gyermekek alaphangmagassága a nemek függvényében óvodáskorban nem mutatott eltérést [ANOVA: $F(1, 4210) = 0,617$; $p = 0,432$], hasonlóan Hasek és munkatársainak (1980) eredményeihez (1. ábra). Az ötéves lányok átlagos alaphangmagassága 261,6 Hz (szórás 49 Hz) volt, míg a fiúk átlagos alaphangja 260,5 Hz (szórás 42 Hz). Az eredmények közelítőleg megfelelnek az angol kutatásokban találtaknak (vö. Lee et al. 1999).

Az általános iskolások esetében a lányok átlagos alaphangmagassága 246,4 Hz (szórás 39 Hz), a fiúké 254,5 Hz (szórás 40 Hz) volt. A két csoport közötti különbség szignifikáns [ANOVA $F(1, 16289) = 161,471$; $p \leq 0,001$]. A kutatásban váratlan eredmény, hogy a fiúk alaphangmagassága a lányokénál magasabbnak bizonyult. Mindez ellentmondani látszik annak a szakirodalmi tapasztalatnak, amely 12 éves korra (serdülőkor) teszi az f_0 elkülönülésének idejét (l. Perry et al. 2001). Deme (2012) 6–7 évesek spontán beszédének vizsgálatakor – a jelen eredményekhez hasonlóan – azt találta, hogy a lányok átlagos alaphangmagassága alacsonyabb. A szerző szerint a lányoknál 6–7 éves korra tehető egy jelentős ugrás az f_0 csökkenésében, amely a fiúknál csak később jelentkezik. Ezzel magyarázható, hogy átmenetileg a lányoknál alacsonyabb alaphangmagasság mérhető (Hacki–Heithmüller 1999, idézi

Deme 2012). A jelen kutatás eredményei alapján ez a tendencia még 10 éves korban is megfigyelhető, bár a csökkenés mértéke kisebb, mint a megelőző életkorban. Valószínűsíthető, hogy 10 éves korra a lányok az alaphangmagasság-változás szakaszának végén, míg a fiúk a változás intenzívebb periódusában vannak. A változás tehát nem lineáris (vö. Hasek et al. 1980). Ennek igazolására azonban további életkori csoportok vizsgálata szükséges. Ugyanakkor az alaphangmagasság eltéréseinek vizsgálatakor a szociokulturális (pszichológiai és társadalmi) vonatkozások sem hagyhatók figyelmen kívül, amelyek a beszédprodukciónak befolyásolhatják (Kohn et al. 2012, idézi Deme 2012), illetve a testalkat is további befolyásoló tényező lehet (Perry et al. 2001).



1. ábra

Az alaphangmagasság eloszlása nemek és életkorok szerint

Deme (2012) hat fonéma realizációit vizsgálta 6–7 évesek spontán beszédében. Az F_1 esetében az [a:] és [o], az F_2 -nél az [ε], [i], az F_1 és F_2 -ben pedig az [e:] esetében talált eltérést a nemek között. A jelen kutatásban az óvodások esetében az F_1 -ben az [ɔ], [e:], az F_1 és F_2 -ben az [o], az F_3 -ban az [ɔ], [a:], [ε] hangok esetében találtam szignifikáns eltérést a nemek között. A statisztikai próba (ANOVA) eredményeit az 1. táblázatban foglaltam össze.

Az [ɔ], [a:], [ε] hangok esetében kimutatott különbségek hasonlóak a külföldi szakirodalomban foglaltakhoz (vö. Busby–Plant 1995). Anyagomban az [ɔ], [a:], [ε] hangok esetében az F_3 is szignifikáns különbséget mutatott a nemek között.

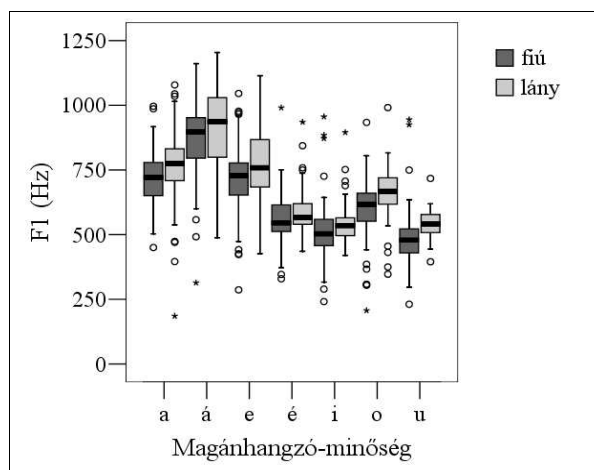
A magánhangzó-minőségek közötti eltérések utalhatnak hangképzésbeli különbségekre, amely fakadhat a toldalékcső méretbeli különbözőségeiből

(Lee et al. 1999). Ezt bizonyíthatja, hogy az eredmények szerint a legtöbb magánhangzó esetében a fiúk formánsértékei alacsonyabbak a lányokénál. A jelen kutatásban a hangképző szervek méretére vonatkozó adat nem állt rendelkezésemre, ezért csak a szakirodalmi tapasztalatoknak megfelelően tudom valószínűsíteni, hogy a fiúk testméretei már ebben az életkorban is nagyobbak a lányokénál.

1. táblázat: A magánhangzók formánsértékeinek nemenkénti eltérése óvodáskorban (a *p* értéke; a szignifikáns eredményeket a szürke szín jelzi)

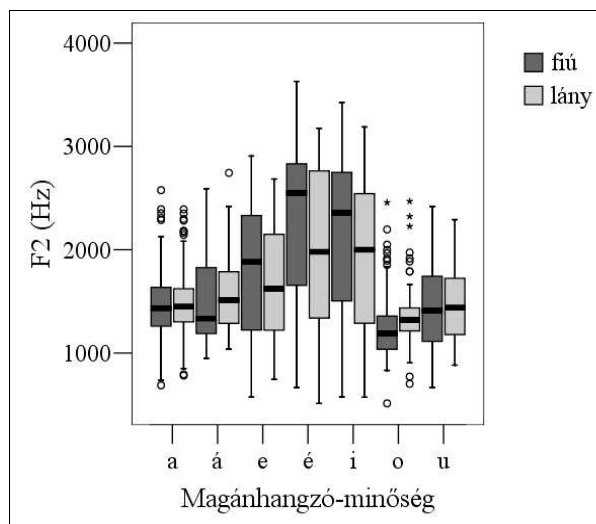
	[ɔ]	[a:]	[ɛ]	[e:]	[i]	[o]	[u]
F ₁	< 0,001	0,170	< 0,001	0,076	0,039	< 0,001	0,072
F ₂	0,694	0,333	0,064	0,095	0,111	0,015	0,458
F ₃	0,001	0,050	< 0,001	0,301	0,014	0,347	0,828

A 2. ábrán látható, hogy az [ɔ], [a:], [ɛ] hangok az F₁ mentén nagy szóródással realizálódtak, ami a nyelvvállásfok artikulációjának variabilitására utal. Ez a változatos megvalósulás az F₂-paraméter esetében nem jelentkezett, egyenletesebbek voltak az értékek (3. ábra). Az adatok közötti átfedések a nemek szerint elkülönített csoportokban mind az F₁, mind az F₂ esetében nagyok voltak, a hangzók minőségének ejtésében ezért valószínűsíthetően kicsi volt a különbség. Az [ɔ], [a:], [ɛ] hangok ejtése a fiúknál zártabb, az [ɛ], [e:], [i] velárisabb volt, hasonlóan Deme (2012) eredményeihez.



2. ábra

A magánhangzók első formánsainak értékei nemek szerint óvodáskorban



3. ábra

A magánhangzók második formánsainak értékei nemek szerint óvodáskorban

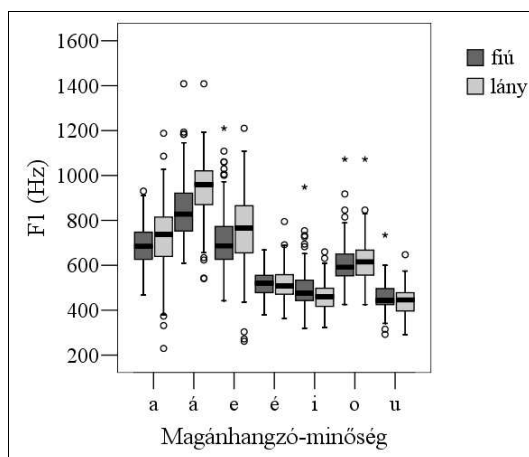
A 10 éves korú gyermekeknel a nemek szerint az F_1 paramétere az alsó és legalsó nyelvállású hangzók ([ɔ], [a:], [ɛ]) és az [i] esetében mutatott szignifikáns eltérést, az F_2 az [a:], [ɛ] és [o] hang esetében, az F_3 pedig az [e:] és [i] hangoknál. A statisztikai próba (ANOVA) eredményeit a 2. táblázatban foglaltam össze.

2. táblázat: A magánhangzók formánsértékeinek nemenkénti eltérése iskoláskorban (a p értéke; a szignifikáns eredményeket a szürke szín jelzi)

	[ɔ]	[a:]	[ɛ]	[e:]	[i]	[o]	[u]
F_1	0,001	< 0,001	< 0,001	0,724	0,001	0,164	0,179
F_2	0,249	< 0,001	0,004	0,198	0,894	0,022	0,880
F_3	0,290	0,031	0,809	0,006	0,002	0,837	0,537

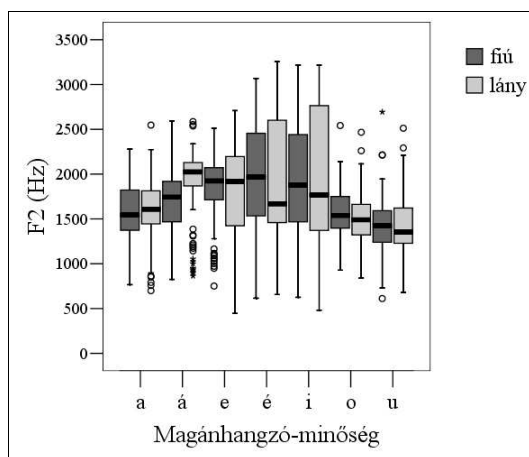
Az eredmények alapján feltételezhető, hogy az alsó és legalsó nyelvállású hangok érzékenyebbek a fizikai növekedéssel bekövetkező változásokra (például toldalékcso méretének változása), hiszen az artikulációt tekintve ezeknél a hangoknál a legnyíltabb a szájüreg és a legtágabb a szájnyílás. Az [ɔ], [a:], [ɛ] hangok F_1 -értékei az óvodásokéhoz hasonlóan nagy szóródást mutatnak, ami változatos artikulációs megvalósulásukra utal (4. ábra). A fiúk ejtése az [ɛ] hangok esetében még zártabb, mint a lányoké. Az F_2 paraméter tekintetében az értékek szóródása kisebb, ami összevetve az 5 éves kori érté-

ekkel, az artikuláció stabilizálódására is utalhat. A variabilitás az F_2 esetében az [ɛ], [e:], [i] hangoknál a legnagyobb (5. ábra).



4. ábra

A magánhangzók első formánsainak értékei nemek szerint iskoláskorban



5. ábra

A magánhangzók második formánsainak értékei nemek szerint iskoláskorban

A percepció teszt eredményei

A kutatásban az óvodások nemét átlagosan 69%-ban, az általános iskolás gyerekeket átlagosan 85,3%-ban osztályozták helyesen a válaszadók (6. áb-

ra). Az 5 évesek esetében a fiúk és lányok felismerési arányának átlaga között nem volt különbség: a lányokat 69%-ban, a fiúkat 69,5%-ban azonosították helyesen, azonban a lányok felismerése a fiúkénál nagyobb szóródást mutatott (a szórás a lányoknál 9,2%, a fiúknál 5,7%), ami a bizonytalanabb észlelést jelzi.

Az iskolások esetében szintén a fiúk felismerése volt pontosabb (95,7%), szemben a lányokéval (75,3%). A lányok azonosítása ebben az életkorban is nagyfokú bizonytalanságot mutatott, a szórás értéke 9,7% volt, szemben a fiúk 1,5%-os értékével.

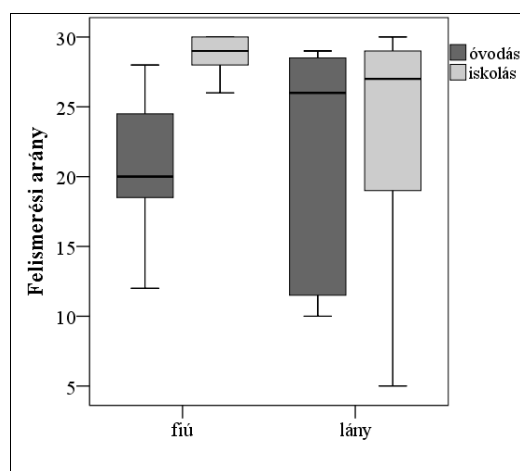
Általánosan elmondható, hogy az életkor előrehaladtával javult a felismerés pontossága a nemekre vonatkozóan. A fiúk felismerése mindkét korcsoportban sikeresebbnek bizonyult.

Elemeztem a felismerési eredményeket az életkor tekintetében is. Az óvodásokat átlagosan 90%-ban, az általános iskolásokat átlagosan 86,4%-ban, tehát közel azonos arányban ismerték fel a percepció tesztet kitöltők. Az észlelésben az életkorra vonatkozólag fontos akusztikai kulcs lehetett az alaphangmagasság, amely az óvodások esetében szignifikánsan magasabb volt, mint az általános iskolásoknál. Mivel a kutatásban csak két korcsoport szerepelt, ezekből az eredményekből messzemenő következtetés nem vonható le az életkorbecslésre vonatkozóan. Tendenciaként azonban megfigyelhető, hogy az óvodáskorú gyermekek életkorának felismerése a humán percepcióban relatíve biztosnak mondható.

A gépi osztályozás eredményei

A kutatásban létrehozott MFCC-alapú SVM-GMM osztályozó az óvodások nemét 48%-ban, az általános iskolásokat 73%-ban osztályozta helyesen. Mivel az óvodások esetében az osztályozás a véletlen határán mozgott, az eredményeket tovább nem elemzem. Az iskolások csoportján belül a lányok osztályozása volt sikeresebb, a gépi címkézés pontossága 85%, a fiúk esetében ez az érték 61,9% volt. A gépi osztályozó eredményei elmaradtak a humán percepciótól. Az a tendencia azonban itt is megfigyelhető, hogy az életkor előrehaladtával javul az osztályozás pontossága. A gépi osztályozás eredménye az életkor tekintetében 47%-os volt, hasonlóan több külföldi kutatáshoz (Feld et al. 2010; Li et al. 2013).

Fontos hangsúlyozni, hogy a gépi osztályozás eredménye a humán percepcióval csak fenntartásokkal hasonlítható össze, mivel nincsenek biztos ismereteink arról, hogy melyek azok a konkrét akusztikai jellemzők, amelyek szerepet kapnak a nem és az életkor azonosításában. A gépi osztályozásban számos jellemző bevonásával próbálják meg modellezni az emberi hallás – ezáltal észlelés – folyamatát (l. például Müller 2006), azonban ezek a jellemzők csak az adott tanítói bázist veszik (és vehetik) alapul, ami az emberi tapasztaláshoz, a szocializációs folyamat során tanult mintákhoz képest relatíve kis „korpuszt” jelent. Ennek is köszönhető, hogy a gépi osztályozás eredményei elmaradnak a humán észleléstől.



6. ábra

Felismerési eredmények a nemek és az életkor függvényében

Következtetések

A jelen kutatásban 5 és 10 éves gyermekek nemének és életkorának becslését vizsgáltam a spontán beszédük alapján. Kutatásom a magyar nyelvben elsőként foglalkozik a gyermekek nemének és életkorának becslésével, illetve a percepció és akusztikum ilyen irányú kapcsolatával.

Feltételeztem, hogy a gyermekek nemének és életkorának azonosítása lehetséges a beszéd alapján. Első hipotézisem igazolódott, az óvodások csoportjában a lányokat 69%-ban, a fiúkat 69,5%-ban azonosították helyesen a válaszadók. Az iskoláskorú gyerekeknél a pontosabb felismerésben az alaphangmagasság és egyes hangoknál a magánhangzóformánsok fontos akusztikai kulcsként működhettek a nemek felismerésében. Az óvodások esetében a magánhangzóformánsok mellett más akusztikai jegyek (például szupraszegmentális térhez köthető jellemzők) befolyásolhatták az észlelést. A kutatás felismerési eredményei hasonlatosak a szakirodalomban foglaltakhoz (vö. Amir et al. 2012). Az a tapasztalat, miszerint a fiatalabb életkorú beszélőket inkább lánynak feltételezik, a jelen kutatásban nem volt igazolható (vö. Traunmüller–Bezooijen 1994).

Második hipotézisem – amely szerint a felismerési pontosság javul az életkor előrehaladtával –, szintén igazolódott, hiszen az iskolás csoportban a lányokat már 75,3%-ban, a fiúkat pedig 95,7%-ban azonosították helyesen a kísérletben részt vevők. Az életkor felismerése mind az óvodások (90%), mind az iskolások (86,4%) esetében sikeresnek bizonyult.

Az akusztikai elemzésből kiderült, hogy az óvodásoknál az alaphangmagasságban a nemek szerint nem adatolható eltérés. A magánhangzó-formánsok elemzésékor az F_1 és F_2 értékei csak bizonyos magánhangzók esetében hordoz-

tak különbséget a nemtől függően. Az alsó és a legalsó nyelvvállású hangok esetében az F_3 értékében eltérést találtam a két csoport között. Az F_3 paraméter ezért a jelen kutatás szerint az alsó nyelvvállású hangoknál hordozhat a beszélő nemére vonatkozó információt. Az F_1 -értékek magas szóródásából arra következtettem, hogy az [ɔ], [a:], [ɛ] hangokat mind az óvodások, mind az iskolások esetében változatos artikulációs megvalósulás jellemezte. Az alsó és legalsó nyelvvállású hangok genderspecifikussága az iskolások csoportjában is megmutatkozott, vagyis ezek érzékenyebbek lehetnek a vokális traktus fizikai növekedés során bekövetkező méretbeli változásaira, mivel képzésüket tekintve a legnyíltabb hangzók.

Az iskolások esetében a nemek között az alaphangmagasságban adatoltam eltérést. A kutatásban a fiúk alaphangja magasabb volt a lányokénál, hasonló jelenséget Deme (2012) is megfigyelt. Ugyanakkor több kutatásban találták azt, hogy a lányoknál 6–10, a fiúknál 8–10 éves kor között megfigyelhető az alaphangmagasság csökkenésének tendenciája (Whiteside–Hodgson 1999). A jelen kutatás eredményeiből az derül ki, hogy ez a folyamat még 10 éves korban is tart. Ezért vannak olyan időpillanatok a növekedésben, amikor ez a nemként eltérő ütemű változás a lányoknál – időlegesen – alacsonyabb alaphangmagasságot eredményez (Deme 2012). A kérdés pontosabb vizsgálatához azonban további kutatások szükségesek. A növekedésbeli, fizikai változások mellett természetesen feltételezhető egyéb tényezők – például szocializációs, tanult hatások – befolyása is.

A nem- és életkorbecslés a beszélőfelismerő rendszerek fejlesztésében is fontos lépés. A kutatásban megalkotott gépi osztályozó eredményei elmaradtak a humán észleléstől. A korpusz és az életkori csoportok (gyerekek) sajátosságai miatt az eredmények a külföldi kutatásokkal csak fenntartásokkal vethetők össze (vö. Li et al. 2013). A gépi osztályozás az iskolások csoportjában a lányok esetében sikeresebb volt, ami a jelen kutatásban a humán percepcióra kapott eredményekkel ellentétes, azonban a külföldi szakirodalmi tapasztalatoknak nem mond ellent (vö. Amir et al. 2012). Bár a humán percepció számára az óvadás és iskolás életkor elkülönítése nem okozott problémát, a gyenge gépi osztályozási eredmény magyarázható az egymáshoz közeli életkorokkal. Több kutatásban találták azt, hogy az életkort célzó gépi felismerés az egymástól távolabbi életkorok esetében sikeresebb (Müller 2006).

Az észlelés és felismerés folyamatában természetesen további jellemzők is szerepet kaphatnak, például a testalkat, illetve szocializációs, kulturális hatások. A kérdés további vizsgálata újabb akusztikai jellemzők bevonásával képzelhető el, ilyenek lehetnek például a beszédhang-minőség jellemzői (jitter, shimmer, időtartam stb.), illetve a szuprasegmentális jegyek (például beszédtempó, beszéd dallam, hangsúly) is. A jelenlegi eredmények a beszélői profilalkotás területén szolgálhatnak adalékokkal, hozzájárulnak a nemek közötti különbségek beszédbeli megjelenésének megismeréséhez, ezáltal a beszéd- és beszélőfelismerő rendszerek fejlesztéséhez.

Irodalom

- Amir, Ofer – Engel, Merav – Shabtai, Esther – Amir, Noam 2012. Identification of children's gender and age by listeners. *Journal of Voice* 26/3. 313–321.
- Balázs Boglárka 1993. Az időskori hangképzés jellemzői. *Beszédkutatás* 1993. 156–165.
- Beke András 2011. Szókezdetek automatikus osztályozása spontán beszédben. *Magyar Nyelvőr* 135/2. 226–241.
- Beke András 2012. Beszélőfelismerés kevert-Gauss modellekkel. In Markó Alexandra (szerk.): *Beszédtudomány. Az anyanyelv-elsajátítástól a zöngékezdési időig*. ELTE BTK–MTA, Budapest. 335–353.
- Beke András – Szaszák György 2009. A svávariációk automatikus felismerése magyar nyelvű spontán beszédben. *Beszédkutatás* 2009. 148–169.
- Boersma, Paul – Weenink, David 2013. *Praat: doing phonetics by computer [Computer program]. Version 5.3*. <http://www.praat.org>. (Letöltés ideje: 2013. júl. 9.)
- Busby, Peter A. – Plant, Geoff 1995. Formant frequency values of vowels produced by preadolescent boys and girls. *Journal of the Acoustical Society of America* 97/4. 2603–2606.
- Childers, Donald G. – Wu, Ke 1991. Gender recognition from speech. Part I: Coarse analysis. *Journal of the Acoustical Society of America* 90/4. 1828–1840.
- Deme Andrea 2012. Óvodások magánhangzóinak akusztikai jellemzői. In, Markó Alexandra szerk.: *Beszédtudomány. Az anyanyelv-elsajátítástól a zöngékezdési időig*. ELTE BTK–MTA Nyelvtudományi Intézet, Budapest. 77–99.
- Diehl, Randy L. – Lindblom, Björn – Hoemeke, Kathryn A. – Fahey, Richard P. 1996. On explaining certain male-female differences in the phonetic realization of vowel categories. *Journal of Phonetics* 24. 187–208.
- Feld, Michael – Burkhardt, Felix – Müller, Christian 2010. Automatic speaker age and gender recognition in the car for tailoring dialog and mobile services. In: *Proceedings of Interspeech 2010*. Chiba. 2834–2837.
- Gocsál Ákos 1998. Életkorbecslés a beszélő hangja alapján. *Beszédkutatás* 1998. 122–134.
- Gorham-Rowan, Mary M. – Laures-Gore, Jacqueline 2006. Acoustic-perceptual correlates of voice quality in elderly man and women. *Journal of Communication Disorders* 39. 171–184.
- Gósy Mária 2001. A testalkat és az életkor becslése a beszéd alapján. *Magyar Nyelvőr* 125. 137–148.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- Harb, Hadi – Chen, Liming 2005. Voice-based gender identification in multimedia applications. *Journal of Intelligent Information Systems - Special issue: Intelligent multimedia applications* 24/2. 179–198.
- Hasek, Carol S. – Singh, Sadanand – Murry, Thomas 1980. Acoustic attributes of preadolescent voices. *Journal of the Acoustical Society of America* 68. 1262–1265.
- Hu, Yakun – Wu, Dapeng – Nucci, Antonio 2012. Pitch-based gender identification with two-stage classification. *Security and Communication Networks* 5/2. 211–225.
- Huber, Jessica L. – Stathopoulos, Elaine T. – Curione, Gina M. – Ash, Theresa A. – Johnson, Kenneth 1999. Formants of children, women, and men: the effects of vocal intensity variation. *Journal of the Acoustical Society of America* 106/3. 1532–1542.
- Huszár Ágnes 2009. *Bevezetés a gendernyelvészetbe*. Tinta Könyvkiadó, Budapest.
- Kassai Ilona 2002. *Fonetika*. Nemzeti Tankönyvkiadó, Budapest.

- Lass, Norman J. – Davis, Margaret 1976. An investigation of speaker height and weight identification. *Journal of the Acoustical Society of America* 60/3. 700–703.
- Lass, Norman J. – Hughes, Karen R. – Bowyer, Melanie D. – Waters, Lucille T. – Bourne, Victoria T. 1976. Speaker sex identification from voiced, whispered, and filtered isolated vowels. *Journal of the Acoustical Society of America* 59/3. 675–678.
- Lee, Sungbok – Potamianos, Alexandros – Narayanan, Shrikanth 1999. Acoustics of children's speech: developmental changes of temporal and spectral parameters. *Journal of the Acoustical Society of America* 105/3. 1455–1468.
- Li, Ming – Han, Kyu J. – Narayanan, Shrikanth 2013. Automatic speaker age and gender recognition using acoustic and prosodic level information fusion. *Computer Speech and Language* 27/1. 151–167.
- Müller, Christian 2006. Automatic recognition of speakers' age and gender on the basis of empirical studies. In *Proceedings of Interspeech 2006*. Pittsburgh. 17–21.
- Perry, Theodore L. – Ohde, Ralph N. – Ashmead, Daniel H. 2001. The acoustic bases for gender identification from children's voices. *Journal of the Acoustical Society of America* 109/6. 2988–2998.
- Ptacek, P. H. – Sander, E. K. 1966. Age recognition from voice. *Journal of Speech and Hearing Research* 9. 273–277.
- Rao, Rajeshwara R. – Nagesh, A. 2011. Source feature based gender identification system using GMM. *International Journal on Computer Science and Engineering* 3/2. 586–593.
- Reynolds, Douglas A. 1995. Speaker identification and verification using Gaussian mixture speaker models. *Speech Communication* 17. 91–108.
- Schwartz, Martin F. 1968. Identification of speaker sex from isolated voiceless fricatives. *Journal of the Acoustical Society of America* 43/5. 1178–1179.
- Traunmüller, Hartmut – van Bezooijen, Renée 1994. The auditory perception of children's age and sex. In: *Proceedings ICSLIP*. 1171–1174. <http://www2.ling.su.se/staff/hartmut/barn.pdf> (Letöltés ideje: 2013.01.03.)
- Vorperian, Hourii K. – Wang, Shubing – Chung, Moo K. – Schimek, E. Michael – Durtschi, Reid B. – Kent, Ray D. 2009. Anatomic development of the oral and pharyngeal portions of the vocal tract: An imaging study. *Journal of the Acoustical Society of America* 125/3. 1666–1678.
- Whiteside, Sandra P. – Hodgson, Carolyn 1999. Acoustic characteristics in 6–10-year-old children's voices: some preliminary findings. *Logopedics Phoniatrics Vocology* 24. 6–13.
- Whiteside, Sandra P. – Hodgson, Carolyn – Tapster, C. 2002. Vocal characteristics in pre-adolescent and adolescent children: A longitudinal study. *Logopedics Phoniatrics Vocology* 27. 12–20.
- Xue, Steve A. – Wing Chi Cheng, Regine–Manwa Ng, Lawrence 2010. Vocal tract dimensional development of adolescents: An acoustic reflection study. *International Journal of Pediatric Otorhinolaryngology* 74. 907–912.

Köszönöm Beke Andrásnak a gépi osztályozásban nyújtott segítségét.

A CÉLZOTT BESZÉDPERCEPCIÓ- ÉS BESZÉDÉRTÉS- FEJLESZTÉS HATÁSA UTÁNKÖVETÉSES VIZSGÁLATBAN

Szántó Anna

Bevezetés

A gyermek anyanyelv-elsajátítására természetesnek tűnő folyamaként tekintünk. A mindennapi életben csak akkor figyelünk föl rá, amikor probléma tapasztalható az elsajátítás menetében. A beszédprodukciónál hibáit, nem megfelelő szintjét könnyebben vesszük észre (például a beszédindulás késése, kórosan hibás artikuláció), míg a percepció folyamatok rejtett működésének zavarai gyakran csak később derülnek ki a kompenzációs stratégiák alkalmazása miatt (Gósy 1994). A gyermek beszédmegértési nehézségei sokszor rejtve maradnak, mivel a beszédhelyzet, s az egyértelműen azonosítható paralingvisztikai tényezők a jó intellektusú gyermeket átsegítik a pillanatnyi problémákon (Macher 2002). A beszédpercepció folyamatának különböző szintjei vannak, amelyek az aktuális értési feladatnak megfelelően különféleképpen aktivizálódnak. A gyermek beszédelsajátítása folyamán szoros kapcsolatban fejlődik az úgynevezett artikulációs és percepció bázis. Az artikulációs bázis biztosítja azt, hogy a gyermek képes legyen anyanyelve beszédhangjait, hangkapcsolatait, szavait hibátlanul kiejteni. A percepció bázis teszi lehetővé, hogy a gyermek képessé váljon a mások által kiejtett beszédhangok, hangsorok felismerésére, a szavak és a mondatok megértésére. A kettő együttműködése eredményezi az anyanyelvi beszédprodukciónak és az anyanyelvi beszédmegértés fejlődését és életkornak megfelelő működését (Gósy 1995/2006). A beszéd percepcióját tekintve az anyanyelv-elsajátítás folyamatában 5-6 éves korra egyértelművé válik a magyar beszédhangok felismerése, függetlenül attól, hogy ki ejti ki azokat, vagy milyen hangkörnyezetben fordulnak elő. A kisiskolás gyermek képes arra, hogy a megfelelő grammatikai, szintaktikai kapcsolatokat felismerje és alkalmazza. Az iskolai tanulás, az írott nyelv elsajátítása gyorsítja ezt a folyamatot, de csak abban az esetben, ha az iskolába lépő gyermek beszédfolyamatainak életkori szinten vannak.

Napjainkban a beszédpercepció folyamatát tekintve a legtöbb kutató a hierarchikus felépítésű modelleket fogadja el. A hierarchikus felépítésű modell szerint a megértési folyamat egymásra épülő szinteken megy végbe az akusztikus jelfeldolgozástól a teljes megértésig. A hierarchikus modellek felépítése nem zárja ki az információ feldolgozásának szeriális és/vagy párhuzamos, illetve moduláris és/vagy interaktív módon történő megvalósulását. Így a hierarchia nem jelent (feltétlenül) egymásutániságot vagy kizárólagos „alulról

felfelé” irányulást. Az egyes részfolyamatok egymással egy időben zajlanak, illetőleg a magasabb szintű folyamatok eredménye is befolyásolja az alacsonyabb szintű működését (Gósy 2005). A teljes percepció folyamatban más-más részfolyamatok működése válik hangsúlyosabbá attól függően, hogy milyen nyelvi egység feldolgozását végezzük éppen (Markó 2007). A beszédészlelés és a beszédmegértés folyamatát számos tanulmány vizsgálta különböző aspektusból (Gósy 2000; Simon 2001; Macher 2002, 2007; Vančóné Kremmer 2002; Erősné et al. 2003; Gósy–Horváth 2006; Imre 2006, 2007; Bóna 2007; Grácsi 2007; Markó 2007; Bartha 2013 stb.).

Miután több tanulmány eredményei is felhívják a figyelmet a beszédészlelés, beszédmegértés elmaradásaira, a fejlesztés, megelőzés jelentőségére, egy saját fejlesztési eljárás kidolgozására került sor. A kutatás célja bemutatni a nagycsoportos óvodások, illetve az első osztályos kisiskolások észlelési, megértési teljesítményének alakulását.

Hipotéziseim szerint a vizsgálati csoportban lévő óvodásoknál, illetve iskolásoknál a fejlesztés hatására javulás lesz tapasztalható (jelentős különbségek az év elején mért adatokhoz képest). A vizsgálati csoportban a fejlesztett gyermekek eredményei jobbak lesznek a korban illesztett, kontrollcsoportban lévő gyermekek eredményeihez képest.

Anyag, módszer, kísérleti személyek

A gyermekek beszédészlelési és beszédmegértési folyamatának felmérését a Gósy Mária által kidolgozott GMP-teszttel végeztem. A GMP sztenderdizált eljárással (Gósy 1995/2006) a magyar gyermekek beszédfeldolgozási folyamata vizsgálható 3–13 éves kor között. Az eredmények értékelését a teszt-csomag életkori sztenderdjei alapján végeztem el. A 2011/2012-es tanév őszi félévében 56 budapesti (28 nagycsoportos óvoda, 28 első osztályos) gyermeket teszteltem a diagnosztika egyes részesztjeivel. Valamennyien ép intellektusú, ép halló gyermekek, egynyelvű fővárosi családokból. A gyermekek hagyományos oktatási rendszerű iskolába járnak. A vizsgált csoport szociológiai szempontból heterogén. A két korcsoportot további két csoportra bontottam, így jöttek létre a vizsgálati, illetve a kontrollcsoportok. A vizsgálati csoportok a diagnosztikát követően hét hónapon át tartó fejlesztésen vettek részt. Miután ekkora létszámba nem készültek még fejlesztési tervek, ezt Gósy Máriával egyeztetve készítettem el a *Beszédpercepció fejlesztő modulok* (Gósy–Imre 2007) című könyv alapján. A kötet öt különböző terület változatos fejlesztő anyagait tartalmazza. Az összetett modulok elsősorban a beszédészlelés fejlesztését szolgálják, de közvetve hozzájárulnak a szókincs bővüléséhez és a nyelvtani tudatosság kialakulásához is. A modulelemek további célzott beszédészlelési fejlesztésre szolgálnak, előkészítik az írott nyelv tanulását. A piramismodulok a beszédészlelés, közvetve a beszédmegértés, illetőleg a munkamemória fejlesztését szolgálják. A szókincsmodulok a gyermekek mentális lexikonának fejlesztését, bővítését teszik lehetővé; közvetve

hozzájárulnak a lexikális aktiválás és asszociációs működés fejlődéséhez is. A lexikonmodulok célja a nyelvhasználat, valamint a spontán beszéd és beszéd-megértés aktív fejlesztése; ezek a szójelentések összefüggéseinek és a grammatikai viszonyok felismerésének javulását eredményezik (Gósy–Imre 2007: 3). A gyerekek minden foglalkozás alkalmával más-más modultípussal foglalkoztak. A fejlesztést a tervek ismertetését, megbeszélését követően a pedagógusok végezték. Az iskolában minden anyanyelvi óra elején vagy végén történt, heti négy alkalommal (5 perc), az óvodásoknál heti öt alkalommal, napközben (15–20 perc). A 2011/2012-es tanév tavaszi félévének a végén, megismételve az év eleji vizsgálatokat, újból felmértem mind az 56 gyermeket.

Jelen tanulmányban a felvett tesztek közül három értékelésére térek ki (GMP12, GMP16, GMP17). A GMP12 rövid állatmesével vizsgálja a gyermek beszédértésének helyes működését. A szöveg szókinccse és szintaktikája megfelel a hatéves gyermek anyanyelv-elsajátítási szintjének. A szövegértési feladat egy mese (rövid, férfi hanggal rögzített szöveg) meghallgatásából, majd tíz megértést ellenőrző kérdésből áll. A mondatértési teszt (GMP16) az adott életkorban elvárt szemantikai sajátosságok ismeretének és szintaktikai/grammatikai struktúrák feldolgozásának szintjéről nyújt felvilágosítást. Anyaga 10 rajzpár, amelyek közül a helyes kép kiválasztása révén mérhető a gyermek mondatértési szintje.

A beszédpercepció egyik fontos mérőköve az a pillanat, amikor a gyermekek képesek lesznek az anyanyelvükben létező egyes hangokat egymástól függetlenül is észlelni, illetve azokat egymással összevetni. A beszédhang-differenciálás vizsgálatához a GMP-diagnosztika (Gósy 2006) 17-es altesztjét használtam, amely 23 egy, két és három szótagú értelmetlen hangsort tartalmaz. A 15 hangorpár egymástól egyetlen képzési jegyben különbözik (például *aszá ~ asá*); 8 pedig teljesen azonos (például *begi ~ begi*). A magánhangzók az időtartam, a mássalhangzók a képzési mód, a képzési hely, a zöngéesség és az időtartam tekintetében különböznek egymástól. A gyermekeknek egyszeri elhangzás után kell eldönteni, hogy egyforma vagy nem egyforma volt-e a hallott két hangsor.

A tesztfelvételhez a szülők minden esetben hozzájárultak. A teszteket a gyermekekkel egyénileg vettem fel, csendes szobában, délelőtt. A tesztfelvétel időtartama átlagosan 25 perc volt. Az adatok statisztikai elemzése az SPSS 13.0 szoftverrel történt.

Eredmények

Óvodások

Miután a nagycsoportos óvodások esetében előfordult, hogy az egyes gyermekek közötti korkülönbség több mint egy év, így az előmérés idején a GMP12-teszt minimálisan elvárt átlagértékét 60%-ban, utóméréskor 70%-ban határoztam meg.

A vizsgálati csoportban az év elején mindössze 7 gyermek teljesítette az elvárható 60%-ot vagy afölötti értéket, ez a kontrollcsoport esetében 10 gyermek. Elemeztem, hogy az óvodások vizsgálati, illetve kontrollcsoportjában az idő előrehaladtával javulnak-e a GMP12- és GMP16-értékek. Az elemzés nem parametrikus Wilcoxon előjeles rangpróbával (összetartozó min-tás) és Monte Carlo szimulációval kiegészítve készült. A GMP12-ben az óvodásoknál a vizsgálati csoportban az év eleji és az év végi eredmények között szignifikáns különbség mutatkozik ($Z = -2,952$, $p = 0,001$), azaz a javulás statisztikailag igazolható. 14 gyermekből 12 gyermeknek javult a teljesítménye, 2 gyermeknek stagnált. A javulás mértéke 10–70% közé tehető. Így például az év elején 20%-on (3 évesektől elvárható) teljesítő gyermek az év végére 90%-on, a 7 évesektől elvárható szinten teljesített. A kontrollcsoport esetében nem találtam szignifikáns különbséget az év eleji, illetve az év végi eredmények között. Ugyan 14 gyermekből 10 gyermeknek javult a teljesítménye, 3 gyermeknek romlott, 1 gyermeknek stagnált. A javulás mértéke 10–50% közé tehető, míg a romlás mind a három gyermek esetében 20%. Év végére a vizsgálati csoportban lévő gyermekek fele teljesít 90–100% között, míg ez a kontrollcsoportban 3 gyermek esetében mondható el.

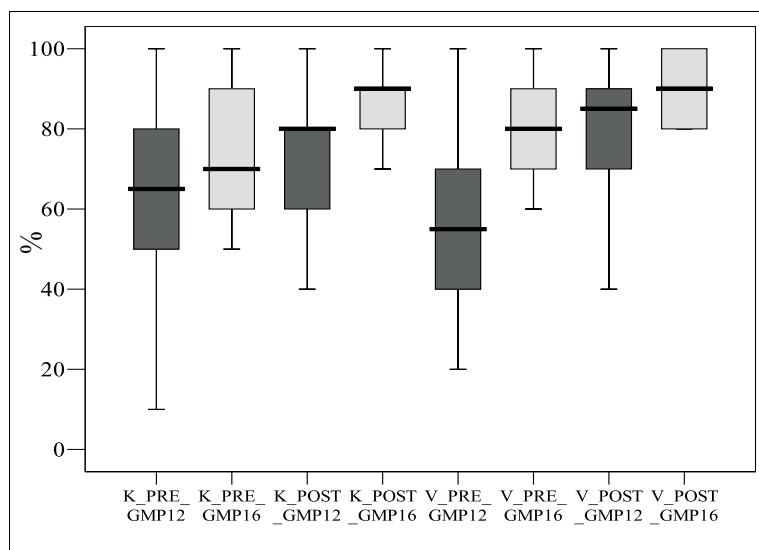
A GMP16-ban a 6 év felett elvárható teljesítmény 90–100%. A gyerekek közötti különbséget figyelembe véve az év elején jónak ítéltm már a 60%-100% közötti eredményt. Mind a két csoport esetében az év eleji, illetve az év végi eredmények között szignifikáns különbség tapasztalható. A vizsgálati csoportban az év végi eredményeket tekintve 8 gyermek teljesítménye javult, míg 6 gyermek teljesítménye stagnált az év elején mértekhez képest ($Z = -2,558$, $p = 0,003$). A javulás mértéke 10–30% közötti. A kontrollcsoportnál 11 gyermeknél tapasztalható javulás (10–30%), 2 gyermek eredménye 10%-kal romlott, míg 1 gyermek eredménye változatlan maradt az év eleji teljesítményhez képest ($Z = -2,623$, $p = 0,005$) (1. ábra).

Megvizsgáltam, hogy a két csoport év eleji és év végi eredményei között a GMP12-t, illetve a GMP16-ot illetően fennáll-e szignifikáns különbség. Noha statisztikai különbség nincs közöttük, az év végi átlagértékek a vizsgálati csoportnál mind a két teszt esetében jobbak.

Iskolások

Az iskolások esetében a 7 évesektől elvárható teljesítmény a GMP12-ben 90-100%-ra tehető. Év elején mind a két csoportból 1-1 gyermek teljesít a korosztálytól elvárható szinten. A vizsgálati csoportban 14 gyermekből az év végére 6 gyermek teljesítménye romlott, a romlás mértéke 10–20% közöttire tehető. 5 gyermek teljesítménye javult, 3 gyermek teljesítménye stagnált. A javulás mértéke 10–50%. Szignifikáns különbséget nem találtam az év eleji és az év végi adatok között. Ellenben a kontrollcsoport eredményeit tekintve az év eleji, illetve az év végi eredményeik között szignifikáns különbség tapasztalható ($Z = -2,949$, $p = 0,001$). Ez az eredmény köszönhető annak is, hogy az utómérést megelőzően egy hónappal anyanyelvi órán feldolgozták a

mesét, ami a szövegértési eredményre pozitív hatással volt. Bár a legjobban és a legrosszabban teljesítő gyermekek közötti teljesítménykülönbség így is 50%.



1. ábra

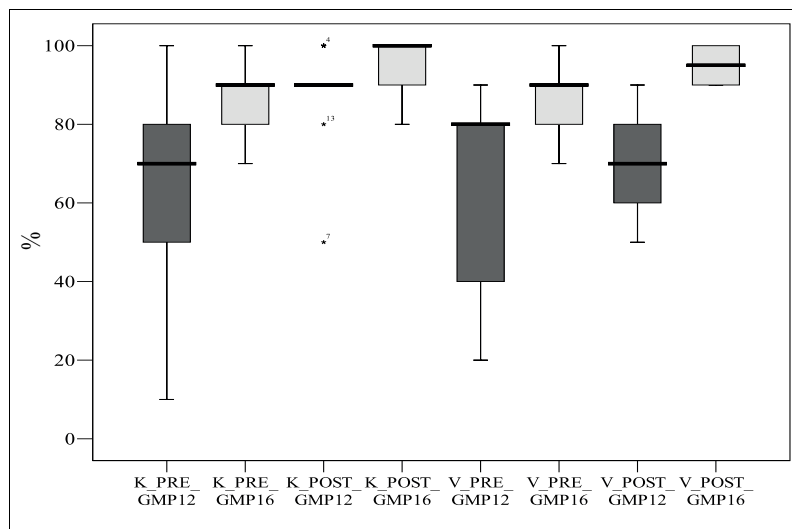
A szöveg- és mondatértési teljesítmény alakulása az óvodásoknál, év elején és év végén (szóródás és medián).

(V=vizsgálati, k=kontroll, pre=előmérés, post= utómérés.)

Iskolásoknál a GMP16-ban elvárható átlagérték 100%, ugyanakkor 90%-os eredményt is elfogadhatónak ítélttem (6 és fél éves korig még a 90%-os eredményt is életkori szintűnek tartjuk). Év végére fejlesztéstől függetlenül szignifikánsan jobb eredmények születtek mind a két csoportban. A javulás mértéke a vizsgálati és a kontrollcsoportban is 10–30%. A mondatértési tesztek alapján megállapítható, hogy mind a két csoport az életkorának megfelelő szinten teljesít. A statisztikai elemzés alapján látható, hogy a két csoport év eleji, év végi eredményei között sem a GMP12-ben, sem a GMP16-ban nem mutatható ki szignifikáns különbség (2. ábra).

A beszédhang-differenciálás egyfelől a fonetikai, másfelől a fonológiai szint működését egészíti ki (Gósy 2006). A beszédhang-differenciálás képessége kiemelt jelentőséggel bír. Ha a beszédhangok megkülönböztetésének képessége nem válik stabilá legkésőbb hétéves korra, ez akadályozza az anyanyelvi szókincs bővülését, nehezíti az írás, helyesírás elsajátítását, de nehézséget okoz az idegen nyelv tanulásában is (Imre 2007). Hétéves korra elvárható, hogy a gyermekek az anyanyelvük beszédhangjait meg tudják kü-

lönböztetni egymástól, függetlenül attól, hogy két hangzó milyen mértékben és milyen akusztikai tulajdonságokban tér el egymástól. Miután az óvodások többsége az év eleji mérésnél még nem töltötte be a 6. életévét, a tesztelésre náluk csak év végén került sor. A fejlesztési tervek tartalmaztak célzott beszédhang-differenciálást fejlesztő feladattípusokat.



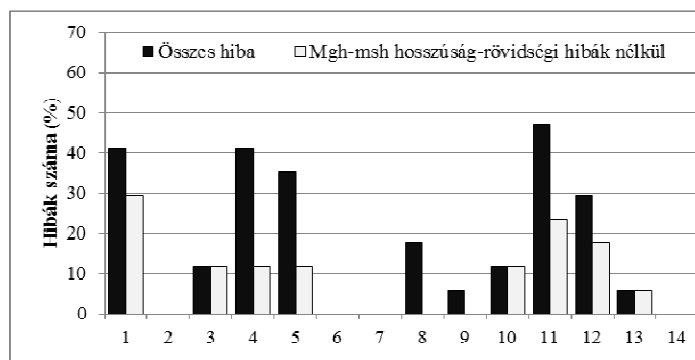
2. ábra

A szöveg- és mondatértési teljesítmény alakulása az iskolásoknál, év elején és év végén (szóródás és medián).

(V=vizsgálati, k=kontroll, pre=előmérés, post= utómérés.)

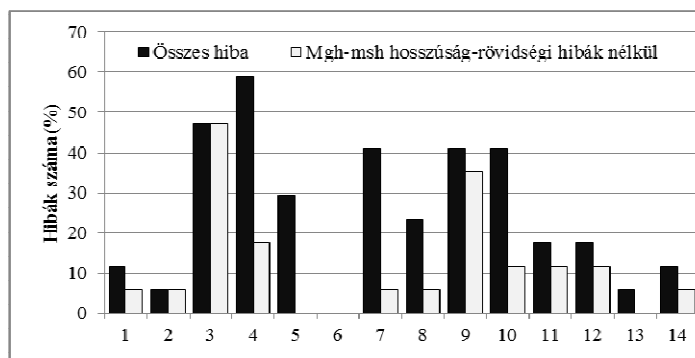
Az óvodás vizsgálati és kontrollcsoport eredményei között nem találtam szignifikáns különbséget (3. és 4. ábra). Hibaelemzést is végeztem, miután az óvodáskorú 6-7 éves gyermekek általában még bizonytalanok a beszédhangok időtartamviszonyainak felismerésében, ami életkori sajátásnak tekinthető. A vizsgálati csoportban 4 gyermek hibátlanul teljesített, további 2 gyermek csak hosszúság-rövidségi hibákat vétett. A kontrollcsoportban összesen 1 gyermek teljesített hibátlanul, míg 2 gyermek az időtartam felismerésében hibázott (3., 4. ábra). Az ábrák x tengelyén a mérésben részt vett gyermekek vannak feltüntetve.

A vizsgálati csoportban átlagosan 17,64%-os a tévesztések száma, azonban ha a hosszúsági és rövidségi tévesztéseket nem számítjuk bele, akkor ez csupán 8,82%. Míg a kontrollcsoportban a tévesztések magasabb arányban vannak jelen, 25,21%-ban (5. ábra)



3. ábra

A GMP17-ben előforduló tévesztések az óvodás vizsgálati csoportban



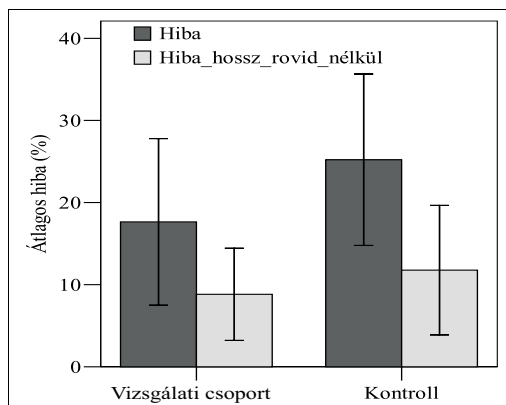
4. ábra

A GMP17-ben előforduló tévesztések az óvodás kontrollcsoportban

Hétéves korra már a magán- és mássalhangzók időtartambeli különbsége sem okozhat gondot. Első osztályosoknak a fekadatot hibátlanul kell megoldaniuk. Megnéztem, hogy a vizsgálati csoport fejlesztés előtt és után milyen javulást ért el, valamint hogy a kapott eredmény szignifikánsnak mondható-e. A fejlesztés előtt átlagosan 31%-os hibát ejtettek, míg utána 4%-osat ($Z = -2,937$, $p = 0,003$). A kontrollcsoportnál 22%-ról csökkent a hibázás aránya 6%-ra ($Z = -2,632$; $p = 0,008$) (6. ábra).

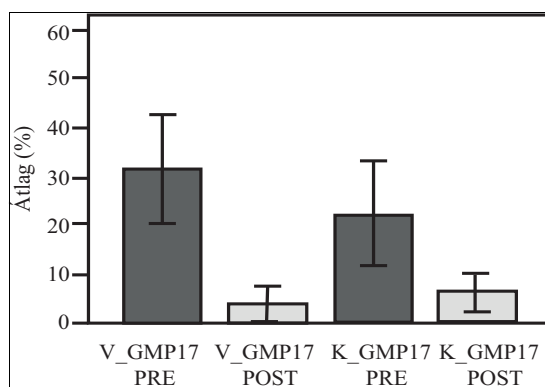
Megnéztem, hogy a fejlesztett és a kontrollcsoport között az év elején, illetve az év végén mért teljesítményekben van-e szignifikáns eltérés a tévesztések számát illetően. Egyik esetben sem találtam szignifikáns különbséget. Az év elején mind a két csoportban csak 1-1 gyermek teljesítette a tesztet hibátlanul. Az év végére a vizsgálati csoportban 9 gyermek teljesítménye

100%-os, 3 gyermeknek egy, 1 gyermeknek kettő, 1 gyermeknek három tévesztése volt. Azok a gyermekek (2), akik az év elején tíz logatompárt ítélték meg tévesen, az év végén hibátlanul teljesítettek. A kontrollcsoportban, év végére 5 hibátlan eredmény született, 4 gyermek egy, 4 gyermek kettő, míg 1 gyermek négy hibát vétett a tesztben. Bár statisztikai különbség nem mutatható ki a két csoport év végi teljesítménye között, az átlagértékeket tekintve a vizsgálati csoportban a gyerekek 64%-a teljesített az életkortól elvárható szinten, míg ez a kontrollcsoport esetében 35% (7. ábra).



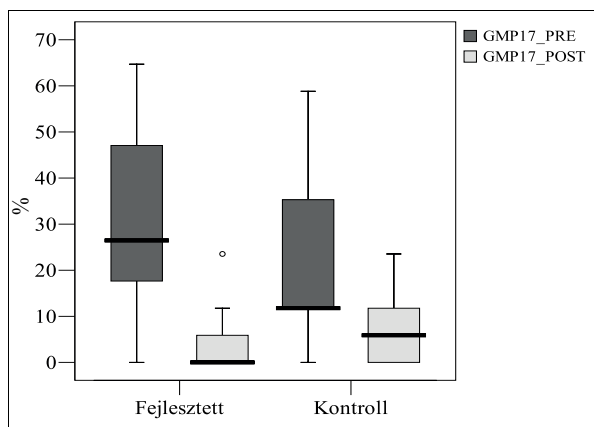
5. ábra

A beszédhang-differenciálási tévesztés átlaga óvodásoknál (%)



6. ábra

A beszédhang-differenciálás tévesztéseinek átlaga év elején és év végén (%)
(V=vizsgálati, k=kontroll, pre=előmérés, post= utómérés.)



7. ábra

A beszédhang-differenciálási tévesztés átlaga és szórása (%).
(Pre=előmérés, post= utómérés.)

Következtetések

A jelen kutatásban a célom nem az volt, hogy a gyermekek beszédfeldolgozásának átlagtól való eltérését jellemezsem. A tanulmány egyik célkitűzése az volt, hogy bemutasson egy olyan alternatív fejlesztési lehetőséget, amely nagyobb csoportban is alkalmazható. Továbbá áttekinteni, hogy célzott fejlesztés hatására milyen módon változnak a tipikus fejlődésű nagycsoportos óvodáskorú, illetve első osztályos kisiskolás gyermekek percepciós mutatói. A korábban már kiértékelt észlelési tesztek (GMP2, GMP3, GMP4, GMP5, GMP10) eredményeihez hasonlóan az itt kapott adatok alapján megállapítható, hogy a fejlesztés hatására óvodáskorban nagymértékű fejlődés tapasztalható, az év eleji egyéni különbségek mértéke jelentősen csökken. Mind a szövegértésben, mind a mondatértésben statisztikailag igazolható a fejlődés. A beszédhang-differenciálás tesztjében a két csoport teljesítménye között nem találtunk statisztikai különbséget, a tévesztések száma a kontrollcsoportban magasabbnak bizonyult. A fejlesztésben részt vett iskolásoknál tendenciásként javultak az eredmények. Bár statisztikailag nem igazolható a két csoport eredményei között lévő különbség, azt lehet mondani, hogy a vizsgálati csoportban önmagukhoz képest a fejlődés mértéke jelentős, az év elején tapasztalható egyéni különbségek év végére csökkennek. A gyengébben teljesítő gyermekek fejlődése nagyobb ütemű, „felzárkóztak” az év elején jobban teljesítő gyermekekhez. Az adatok arra utalnak, hogy az első osztályban nagy szükség van még a fejlesztésre, lehetőség szerint hosszabb időtartamban. Érdeemes áttekinteni a vizsgálati (óvodás) csoport év végi eredményeit, illetve az iskolások év eleji eredményeit (azonos korcsoport). Bár nem ugya-

nazon gyermekekről van szó, de jól látható, hogy valamennyi teszt értékeit tekintve kimutatható különbség áll fenn az óvodások javára. Ez bizonyítékkul szolgálhat arra, hogy miután óvodáskorban több idő, figyelem fordítható a fejlesztésre, a javulás mértéke nagyobb lehet. Ha a fejlesztés jelentősége mellett érvelünk, azt sem árt figyelmen kívül hagyni, hogy a vizsgálatban részt vett gyermekek tipikus fejlődésűek. Jól ismert tény, hogy az anyanyelv-elsajátítás individuális különbségei a tipikus fejlődésűek között is jelentősek (például Smith–Kenney 1998). Továbbá azt sem szabad elfelejteni, hogy a hagyományos oktatási rendszerű iskolába bekerülnek a tipikus fejlődéstől eltérő gyermekek is, akiknek még nagyobb szüksége lenne a fejlesztésre. Korábbi vizsgálatok eredményei (Kocsis 1996; Macher 2002, 2007; Erősné et al. 2003; Bóna 2007 stb.) és az itt kapott adatok alapján is célszerű lenne, ha az iskolaérettségi vizsgálatok során a jó erőnlét, megfelelő fizikai adottság, mozgáskoordináció, látás, artikuláció, hallás szűrése mellett minden esetben elvégeznék a beszédészlelés és beszédmegértés vizsgálatát is. Amennyiben pedig szükséges, az elmaradt területeket, következetes fejlesztő gyakorlással korrigálják. Fontos lenne, mert sok kisgyermek kezd el iskolai tanulmányait az elvárhatót el nem érő beszédészlelési és beszédmegértési szinttel. Később sokszor arra is fény derül a gyermek általános iskolás tanulmányai végétével, hogy az addigra kialakult tanulási nehézség gyökerét bizony a percepció elmaradásokra lehet visszavezetni, mert ha „csak” e terület működése csökkentebb az alsóbb osztályokban, a nem szakember számára a probléma rejtve marad, jelzés nem érkezik sem a szülőktől, sem a pedagógustól (Erősné et al. 2003). A percepció zavarok korai felismerése meghatározó a gyermek egész fejlődésére, különösen az anyanyelv-elsajátítására és a tanulási folyamatokra. A fejlesztés hatékonysága növelhető lenne, ha a diagnosztikát követően egyénre szabottan, célzottan az elmaradt területekre koncentrálna készülnének a fejlesztési tervek. Összegzésként megfogalmazható, hogy a fejlesztés nem volt hiábavaló, szükséges további lépéseket tenni annak érdekében, hogy minél szélesebb körben fordítsanak figyelmet a beszédészlelés és a beszédmegértés folyamataira, különösképpen a fejlesztésre.

Irodalom

- Bartha Krisztina 2013. A beszédfeldolgozás jellemzői magyar domináns kétnyelvű gyermekeknél. *Beszédkutató 2013*. 233–246.
- Bóna Judit 2007. A fonológiai és a szeriális észlelés fejlődése 4–10 éves korban. In: Gósy Mária (szerk.): *Beszédészlelési és beszédmegértési zavarok az anyanyelv-elsajátításban*. Nikol KKt., Budapest. 262–270.
- Erősné Brunner Katalin – Csabay Katalin – Győry Gáborné 2003. Beszédészlelés, beszédmegértés, szövegészlelés. *Gyógypedagógiai Szemle* 2003/2. 88–96.
- Gósy Mária 1994. A mondatértés és a szövegértés összefüggései. *Beszédkutató* '94. 94–120.
- Gósy Mária 1995/2006. *GMP-diagnosztika*. Nikol GMK, Budapest.

- Gósy Mária 2000. *A hallástól a tanuláshoz*. Nikol Kkt., Budapest.
- Gósy Mária 2005. *Pszicholingvisztika*. Osiris Kiadó, Budapest.
- Gósy Mária 2006. A beszédhangok megkülönböztetésének fejlődése. *Beszédkutató 2006*. 147–159.
- Gósy Mária – Horváth Viktória 2006. A percepció folyamatok összefüggései hatéveseknél. *Alkalmazott Nyelvtudomány* IV/1–2. 25–42.
- Gósy Mária – Imre Angéla 2007. *Beszédpercepció fejlesztő modulok*. Nikol Kkt., Budapest.
- Grácz Tekla Etelka 2007. Diszlexiás és tipikus fejlődésű gyermekek beszédfeldolgozásának vizsgálata. In Gósy Mária (szerk.): *Beszédészlelési és beszédmegértési zavarok az anyanyelv-elsajátításban*. Nikol Kkt., Budapest. 202–213.
- Imre Angéla 2006. Az olvasászavar és a beszédfeldolgozási folyamatok összefüggései. *Beszédkutató 2006*. 160–171.
- Imre Angéla 2007. A beszédmegértés és az olvasás összefüggése. In Gósy Mária (szerk.): *Beszédészlelési és beszédmegértési zavarok az anyanyelv-elsajátításban*. Nikol Kkt., Budapest. 164–183.
- Kocsis Judit 1996. Óvodáskorú gyermekek beszédészlelési és beszédmegértési teljesítményének alakulása. In Gósy Mária (szerk.): *Gyermekekori beszédészlelési és beszédmegértési zavarok*. Nikol Kkt., Budapest. 100–121.
- Macher Mónika 2002. Cigány gyermekek beszédészlelésének és beszédmegértésének vizsgálata. *Beszédkutató 2002*. 118–130.
- Macher Mónika 2007. Tanulásban akadályozott gyerekek beszédészlelési és beszédmegértési vizsgálata. In Gósy Mária (szerk.): *Beszédészlelési és beszédmegértési zavarok az anyanyelv-elsajátításban*. Nikol Kkt., Budapest. 247–260.
- Markó Alexandra 2007. A mondat- és szövegértés jellemzői és összefüggése 6-9 éves korban. In Gósy Mária (szerk.): *Beszédészlelési és beszédmegértési zavarok az anyanyelv-elsajátításban*. Nikol Kkt., Budapest. 285–300.
- Simon Orsolya 2001. Ötödik és hatodik osztályosok anyanyelvi szövegértési mutatói. *Iskolakultúra* 2001/11. 92–100.
- Smith, Bruce L. – Kenney, Mary Kay 1998. An assessment of several acoustic parameters in children's production development: longitudinal data. *Journal of Phonetics* 26. 95–108.
- Vančóné Kremmer Ildikó 2002. A beszédészlelés és a beszédmegértés vizsgálata magyar–szlovák kétnyelvű gyermekeknél. In: Lanstyák István – Simon Szabolcs (szerk.): *Tanulmányok a kétnyelvűségről*. Kalligram Könyvkiadó, Pozsony. 71–94.

MEGAKADÁSJELENSÉGEK AZ ÉLETKOR, A NEM ÉS A BESZÉDTÍPUS FÜGGVÉNYÉBEN

Bóna Judit

Bevezetés

A megakadásjelenségek „a spontán beszéd artikulációs, percepciók folyamatosságát megakasztó, illetve a köznyelvi normának ellentmondó különféle jelenségek” (Gósy et al. 2009: 258), amelyek mintegy „ablakok” a beszédtervezés rejtett folyamataira (Goldman-Eisler 1958; Fromkin 1973; Levelt 1989). Két nagy típusuk van: a beszélő bizonytalanságából származó megakadások (bizonyos néma szünetek, hezitálás, nyújtás, ismétlés, újraindítás, töltelékszók, szünet a szóban) és a tervezés és kivitelezés összehangolatlan-ságára visszavezethető hibák (Gósy 2005). Ez utóbbiak a beszédtervezés bármely szintjén létrejöhetnek (Levelt 1989). A mentális lexikon aktiválásához köthető hibák a téves szó, a téves kezdés és a „nyelvem hegyén van” jelenség (TOT); a nyelvi átalakításhoz köthető hibák a grammatikai hiba és a kontamináció; az artikulációs tervezés hibái a sorrendiségi hibák (anticipáció, perszeveráció és metatézis); az artikulációs kivitelezés hibája pedig az egyszerű nyelvbotlás. Vannak olyan jelenségek is, amelyek a felszíni elemzés alapján több típusba is sorolhatók (Gósy et al. 2009).

A megakadások megjelenésének gyakoriságát számos tényező befolyásolja, például: egyéni jellemzők (Gósy 2005), a beszéd- és artikulációs tempó (Bóna 2006), az életkor, a nem és a beszéd típus. A jelen tanulmányban ez utóbbi hárommal foglalkozunk részletesebben.

Az életkor előrehaladtával számos változás történik a beszédben, ami hormonális, fiziológiai, pszichológiai és kognitív változásokra vezethető vissza (Ulatowska et al. 1986; Humes 1996; Ramig et al. 2001; Hnath-Chisolm et al. 2003; Xue-Hao 2003). Az időskori beszéd egyik legszembetűnőbb jellemzője a szótalálási nehézség (Burke et al. 1991; Kemper 1992; Horváth 2005), de általában véve a megakadások gyakorisága is mutathat életkori sajátosságokat. Az időskori beszédben jelentkező megakadásokkal viszonylag kevés tanulmány foglalkozik, és a gyakoriságra vonatkozó adatokban nincs egyetértés a szerzők között. A fiatalok és az idősek beszédprodukciónak az összevetése során egyes szerzők arra az eredményre jutottak, hogy nincs különbség a megakadások gyakoriságában a két életkori csoport között (Duchin-Mysak 1987; Leeper-Culatta 1995); míg más kutatások szerint az idősek beszédprodukciónak több a megakadásjelenség (Yairi-Clifton 1972). Saját vizsgálatunkban azt találtuk, hogy a megakadás típusától is függ,

hogy a két életkori csoport között van-e különbség a gyakoriságban (Bóna 2013b). Egy másik kutatásunkban idősödők, idősek és százévesek megakadásait elemeztük a gyakoriság tekintetében. A bizonytalansági megakadások az életkor előrehaladtával tendenciaszerűen csökkentek, a hibák gyakorisága azonban nőtt (Bóna 2012).

Több tanulmányban azt találták, hogy a férfiak gyakrabban produkálnak megakadást, mint a nők. Shriberg (2001) a fő különbséget a hezitálások gyakoriságában adotta, az általa vizsgált korpuszban a férfiak több kitöltött szünetet produkáltak, mint a nők. Továbbá megállapította, hogy ha a társalgás során a hallgató férfi, akkor a beszélő (függetlenül attól, hogy nő vagy férfi), több megakadást produkál, mint női beszédpartner esetében. Az említett kutatásban a szociolingvisztikai faktorokat nem kontrollálták (a nők és a férfiak iskolázottsága eltérő volt). Kontrollált körülmények között is hasonló eredményekre jutott Bortfeld et al. (2001), mint Shriberg (2001). A Bortfeld és munkatársai (2001) által vizsgált társalgásokban a férfiak több megakadást produkáltak, mint a nők. A férfiak beszédprodukciónak szignifikánsan gyakrabban fordultak elő hezitálások és ismétlések, illetve tendenciaszerűen gyakoribbak voltak az újraindítások is. Egy másik kutatásban (Branigan et al. 1999) azt igazolták, hogy a nők csak akkor produkálnak kevesebb megakadást, mint a férfiak, ha láthatják a beszédpartnerüket. Ebben a kutatásban a beszédpartner neme nem befolyásolta a megakadások előfordulását.

A magyarra vonatkozóan Gósy (2003) megállapította, hogy a férfiak (monologikus) beszédében több a bizonytalanságra visszavezethető megakadás, mint a nőknél, a hiba típusú jelenségek előfordulásában azonban nincs különbség a két nem között. Hasonló eredményekre jutott Horváth (2007) dialogikus felvételek elemzése során. Az eredményei szerint a férfiak nemcsak több megakadást produkáltak, mint a nők, de voltak tipikusan rájuk jellemző megakadástípusok is, és a férfiak beszédprodukciónak többféle megakadástípus jellemezte, mint a nőké. Huszár több írásában (2009, 2011) is megemlíti, hogy a férfiak többet hezitálnak, mint a nők. Hasonló eredményre jutott Gocsál (2001) is. A hezitálások gyakoriságával kapcsolatosan azonban más kutatásokban ellentmondó eredmények olvashatók. Horváth (2010) fiatal nők és férfiak beszédében vetette össze a hezitálások sajátosságait. Az eredményei szerint nem volt szignifikáns különbség a hezitálások percenkénti előfordulásában a két nem között. Egy vizsgálatban (Gósy et al. 2013) 48 adatközlő többórnyi beszédprodukciónak vizsgálták a leggyakoribb előfordulású ö-féle hezitálást. Az eredmények azt mutatták, hogy kilencéves korban még nincs különbség a lányok és a fiúk között a svászerű hezitálások percenkénti előfordulásainak a számában, míg a fiatal felnőtteknél és az időseknél a nők beszédében gyakrabban fordul elő hezitálás ugyanazon időegység alatt. Ugyanakkor a férfiak hezitálásainak az időtartamai szignifikánsan hosszabbak voltak, mint a nőknél mérték.

A beszéd típus is nagymértékben meghatározza a megakadások előfordulását, mivel a különféle beszédhelyzetek eltérő beszédprodukciós stratégiát igényelnek (Gósy 2005). Spontán beszédnek nevezzük mindazon beszédprodukciókat, amelyek során a beszédtervezés és -kivitelezés egyszerre zajlik (Gósy 2005). Ugyanakkor a spontán beszéd többféle beszédhelyzetben, beszédfeladatok során létrehozható, ezek a helyzetek gyakran eltérő mentális erőfeszítést, figyelmet, önmonitorozást igényelnek. A spontán narratívák során a beszélők általában szabadon beszélhetnek, ők választhatják meg a beszéd tartalmát, a szavakat, a grammatikai szerkezeteket, és annyi idő áll a rendelkezésükre, amennyit csak akarnak. Az irányított spontán beszéd létrehozása (pl. képsor elmesélése, játékszabály ismertetése, hallott vagy olvasott szövegek tartalmának visszamondása) már nagyobb mentális erőfeszítést igényel a beszélőtől. Például a hallott szövegek tartalmának összegzése relatíve nehéz, a narratívánál nagyobb mentális erőfeszítést igénylő feladat; a sikerességét több tényező befolyásolja: a figyelmi folyamatok, a munkamemória, a beszédfeldolgozás és a narratív kompetencia (Ulatowska et al. 1986; Juncos-Rabadán et al. 2005). Az adatközlő ilyenkor nem válogathat szabadon a nyelvi elemek közül, hiszen egy más által megírt szöveg szerkezetét, szókincsét kell felidéznie (Gósy 2010). A társalgás is különböző beszédstratégiákat kíván meg, mint egy monologikus beszéd típus. Ez egyrészt egyfajta „verseny”, amely a megszólalás jogáért alakul ki a beszélők között, másrészt a többiek megszólalásai időt adnak a beszélőnek a beszédtervezésre (Markó 2005). A beszéd lépések időtartama a társalgásokban jellemzően rövid, vagy legalábbis sokkal rövidebb, mint az interjú során rögzített narratíváké. A különféle beszédhelyzetekre jellemző tervezésbeli különbségekről a beszédprodukciók temporális sajátosságai is árulkodnak (Markó 2005; Bóna 2013a).

A magyar nyelvre vonatkozóan a megakadások gyakoriságát a spontán narratívák mellett elemezték már térképmódszerrel rögzített dialógusokban (Horváth 2004), nyelvi játék közben (Szabó 2004), különféle monologikus beszéd típusokban (Markó 2004), szinkrontolmácsolásban (Bakti 2009), szónoki beszédben (Laczkó 2010), prezentációk során (Pethő 2013), illetve tartalomösszegzéskor (Bóna 2013b). Ugyanakkor nem készült még olyan átfogó elemzés a megakadások gyakoriságáról, amely relatíve sok adatközlővel többféle beszéd típusban elemezte volna az életkor és a nem hatását a megakadásjelenségek előfordulására.

A jelen kutatás célja annak elemzése, hogy milyen különbségek vannak az idős és a fiatal nők és férfiak megakadásainak gyakoriságában spontán narratíva, tartalomösszegzés és többszereplős társalgás során. Továbbá kérdésünk az, hogy milyen megakadástípusok jellemzik a különböző életkorú beszélőket, a két nemet, illetve a beszéd típusokat.

Hipotéziseink szerint 1. a nagyobb mentális erőfeszítést igénylő beszéd típusban (tartalomösszegzés) megnő a megakadások gyakorisága, míg a versenyhelyzetet jelentő, ugyanakkor a tervezési folyamatokra több időt hagyó

társalgásban csökken a gyakoriság a spontán narratívákhoz képest. 2. Az életkor is befolyásolja a gyakoriságot (az időseknél feltehetően több megakadás lesz mindegyik beszéd típusban), illetve a megakadások típusát is: az időseknél a hibák lesznek gyakoribbak, a fiataloknál pedig a bizonytalanságra utaló megakadások. 3. Feltételezzük továbbá, hogy mindegyik beszéd típusban különbségeket találunk a nők és a férfiak beszéde között is; lesznek olyan megakadástípusok, amelyek inkább a nőkre, illetve inkább a férfiakra jellemzők.

Kísérleti személyek, anyag, módszer

A kutatáshoz 40 adatközlő hangfelvételeit választottuk ki a BEA-ból (Beszélt nyelvi adatbázis, vö. Gósy et al. 2012). Az adatközlők két életkori csoportból kerültek ki: 20–30 év közötti fiatalok és 66–90 év közötti idősök. Mindegyik életkori csoportban 10 nő és 10 férfi beszédprodukciónak elemeztük. Az adatközlők egynyelvű, magyar anyanyelvű, köznyelvet beszélők voltak, ismert mentális és hallási problémájuk nem volt. Figyeltünk arra is, hogy az iskolázottságuk is egyforma legyen minden csoportban; mindegyik adatközlőnek legalább középfokú végzettsége volt.

A felvételekből három részletet elemeztünk: 1. interjúhelyzetben rögzített narratívát (minden adatközlőtől mintegy 5 percet), 2. tartalomösszegzést (ennek az időtartama beszélőfüggő volt), 3. háromszereplős társalgást (adatközlőnként 4,5-5 percet, amely több beszéd lépésből állt össze). Összesen 8,6 órányi beszédet elemeztünk.

A narratíva során a felvételvezető a munkájáról, családjáról, hobbijáról kérdezte az adatközlőket, de csak akkor szólalt meg, és tett fel új kérdést, amikor az adatközlőnek már nem volt mit mondania. A tartalomösszegzés során két rövid szöveg (egy tudományos ismeretterjesztő és egy történelmi anekdota) meghallgatása, majd a lehető legpontosabb visszamesélése volt a feladat. A szövegek lexikai és grammatikai komplexitása átlagos volt, a megértésük semmiféle speciális tudást nem igényelt. A lejátszott ismeretterjesztő szöveg időtartama 97 s volt, összesen 174 szóból állt. A történelmi anekdota időtartama 125 s volt, összesen 270 szóból állt. Az elemzett társalgások háromszereplősök voltak, az adatközlő mellett a felvételvezető és egy kollégája vettek részt a beszélgetésben.

A beszédprodukciónak elemeztük a megakadásjelenségek gyakoriságát, amelyet a szószámhoz viszonyítva határoztunk meg: azaz megadtuk, hány szavanként fordult elő egy megakadás. A megakadások gyakorisága többféleképpen megadható: egyrészt kiszámítható az adott csoportra jellemző átlag az egyéni különbségek figyelembe vétele nélkül. Ilyenkor az adott csoportba tartozó adatközlők összes szószámát osztjuk el az összes megakadás számával. Másrészt kiszámíthatjuk egyénenként is a megakadásgyakoriságot. Ekkor figyelembe kell vennünk azt is, hogy vannak olyan megakadástípusok, amelyek nem minden beszélőnél fordulnak elő. A tanulmányunkban mindkét mutatót kiszámítottuk. Meghatároztuk, hogy általában az adott beszéd típus-

ban mi jellemzi az adott beszélőket, másrészt elemeztük az egyéni különbségeket is.

Ezután megvizsgáltuk a megakadástípusok előfordulását is. A hibajelenségek tipizálásához Gósy et al. (2009) kategóriarendszerét használtuk fel.

A bizonytalansági megakadások között nem vettük figyelembe a néma szüneteket, mivel a spontán beszédben nem dönthető el egyértelműen, hogy az adott szünet megakadás-e, vagy más funkcióban (például levegővétel miatt, esetleg tagolásként) szerepel-e (Gósy 2000). A bizonytalansági megakadásokat a következő típusokba soroltuk:

1. hezitálás: a hazai szakirodalmi hagyományok alapján a kitöltött szünet tartozik ebbe a kategóriába (a nemzetközi szakirodalomban gyakran hezitálásnak tartott töltelékszókat külön elemeztük), a magyarban általában *ö*, *m*, *öm*, *öh* formában fordul elő (Horváth 2010);
2. nyújtás: egyes beszédhangok időtartamának a környezetéhez viszonyított perceptuálisan érzékelhető meghosszabbítása (Bóna 2007);
3. ismétlés: egy teljesen kiejtett szó megismétlése; amikor az ismétlés egyértelműen beszédtervezési nehézségekre, és nem pragmatikai, stilisztikai szerepre utal;
4. újraindítás: amikor egy részben kiejtett szót ugyanazon szó teljes kiejtése követ (Gósy 2005; Gyarmathy 2009);
5. töltelékszó: azok a szavak vagy szókapcsolatok (pl. *tehát*, *akkor*, *hát*, *ilyen*, vagy *nem tudom*), amelyek a közlésbe tartalmilag nem illeszkednek, funkciójukat tekintve pedig a diszharmonia egyfajta feloldását segítik (Gósy–Horváth 2009). Bár van olyan osztályozás, amely egy kategóriába sorolja ezeket a kifejezéseket a hezitálással, a szakirodalmi adatok szerint az életkor kapcsán érdemes a töltelékszók gyakoriságát a kitöltött szünettől külön elemezni (Kemper et al. 1990);
6. szünet a szóban: a szó belsejében tartott szünet, amely különböző szintű tervezési nehézségre vagy önmonitorozásra utal a megjelenés helyétől függően.

A hiba típusú megakadásokat a következőképp tipizáltuk (Gósy et al. 2009):

1. grammatikai hiba: a mai köznyelvi, kodifikált grammatikának ellentmondó jelenség (pl. *és aztán összekeveredjük a két gyilkosság miatt*);
2. kontamináció: nyelvi jelek vegyülése; létrejöhet a szavak és a szerkezetek szintjén is (pl. *meg kell mondjam, nekem tiszta a lelkiifurdalásom* 'tiszt a lelkiismeretem × nincs lelkiifurdalásom');
3. téves szó: hibás aktiválás a mentális lexikonból, a teljes szó kiejtése megtörténik (pl. *Járt ma itt a bádogos, mármint a kőműves?*);
4. téves kezdés: hibás aktiválás a mentális lexikonból, a szó egy részének kiejtése történik csak meg (pl. *Felhívtad a pa a mamádat?*);

5. „nyelvem hegyén van” jelenség: átmeneti szótalálási nehézség; a beszélő tudja aktiválni a szó jelentését és morfológiai struktúráját, de a fonológiai/fonetikai szerkezet kiejtése gátolt (pl. *annál az ízénél, a na, ízénél, munkásnál szerelőnél van a gépiünk*);
6. anticipáció: az a jelenség, amikor a közlés egy későbbi eleme korábban jelenik meg a kiejtésben (pl. *Az még egy fülön fejezet, hogy mit kezdünk*);
7. perszeveráció: az a jelenség, amikor a közlés egy korábbi eleme hat az időben később következő elem artikulációs tervezésére (pl. *a politikai pártok patája után 'csatája után'*);
8. metatézis: egy közlés szavainak vagy egy hangsor beszédhangjainak a felcserélődése (pl. *Kucsd be, kérlek! 'csukd be'*);
9. egyszerű nyelvbottlás: az aktuálisan létrehozott artikulációs mozdulat „megbotlik” (pl. *Ahhoz, hogy most tapasztalatot gyűjtenek gyűjtse-nek...*).

A gyakorisági paramétereken statisztikai elemzést végeztünk (MANOVA, ANOVA, Tukey post hoc teszt) az SPSS 13.0 szoftverrel 95%-os konfidenciaintervallumban.

Eredmények

Összesen 6277 megakadást elemeztünk. A csoportokra jellemző, az egyéni különbségeket figyelmen kívül hagyó gyakorisági értékek a következőképp alakultak: a narratívában a fiataloknál 7,9 szavanként, az időseknél 9,2 szavanként; a tartalom-visszamondáskor a fiataloknál 5,2 szavanként, az időseknél 6,6 szavanként; a társalgásban a fiataloknál 12,1 szavanként, az időseknél 14,3 szavanként adatoltunk megakadást.

Ezután kiszámítottuk egyénenként is a gyakorisági értékeket; az összes megakadásjelenség életkor és beszédtypus szerinti szószámra vetített gyakorisága az 1. táblázatban olvasható. Az átlagokat tekintve megfigyelhető, hogy mindegyik beszédtypusban ritkábban fordultak elő megakadások az időseknél, mint a fiataloknál.

A statisztikai elemzés szerint a legnagyobb hatása a beszédtypusnak volt a megakadások gyakoriságára: $F(2, 119) = 25,482$; $p \leq 0,001$; $\eta^2 = 0,321$. Szignifikáns különbséget találtunk az életkor és a nem együttes hatásában [$F(1, 119) = 5,843$; $p = 0,017$; $\eta^2 = 0,051$], illetve a beszédtypus és a nem együttese is statisztikai különbséget mutatott a gyakoriságban [$F(2, 118) = 3,667$; $p = 0,029$; $\eta^2 = 0,064$].

A férfiak és a nők között csak a fiatalok narratíváiban találtunk szignifikáns különbséget [$F(1, 18) = 4,698$; $p = 0,044$], a többi esetben nem volt statisztikailag kimutatható különbség a két nem megakadásainak gyakorisága (a beszéd folyamatossága) között (1. ábra). Ennek oka a nagy egyéni különbségekben (az adatok szórásában) lehet, mivel az átlagértékek gyakran szinte azonosak voltak a két nemnél ugyanazon életkorban és beszédtypusban. A narratívákban a fiatal

nőknél átlagosan 10,9 szavanként (a szórás: 4,6), a fiatal férfiaknál 7,2 szavanként (a szórás: 2,9); az idős nőknél átlagosan 9,8 szavanként (a szórás: 4,1), az idős férfiaknál 11,9 szavanként (a szórás: 7,1) fordult elő megakadás. A tartalomösszegzésekben a fiatal nőknél átlagosan 5,8 szavanként (a szórás: 1,4), a fiatal férfiaknál 5,8 szavanként (a szórás: 2,8); az idős nőknél átlagosan 5,9 szavanként (a szórás: 2,3), az idős férfiaknál 10,9 szavanként (a szórás: 4,4) adatoltunk megakadást. A társalgásokban a fiatal nőknél átlagosan 17,6 szavanként (a szórás: 10,0), a fiatal férfiaknál 12,4 szavanként (a szórás: 5,8); az idős nőknél átlagosan 16,6 szavanként (a szórás: 6,4), az idős férfiaknál 14,3 szavanként (a szórás: 4,3) jelent meg megakadás.

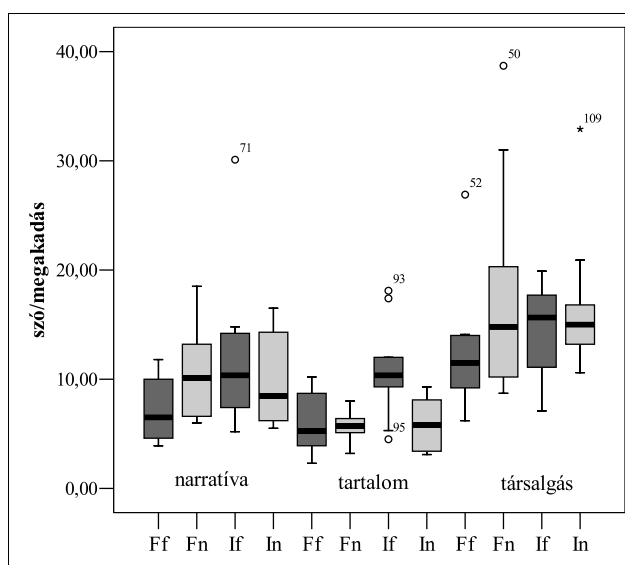
1. táblázat: A beszéd folyamatossága (az összes megakadás gyakorisága) az életkor és a beszéd típus függvényében (szó/megakadás, minél magasabb az érték, annál folyamatosabb a beszéd)

Beszéd típus	Átlag	Átlagos eltérés	Minimum–maximum
Fiatalok			
Narratíva	9,1	4,2	3,9–18,5
Visszamondás	5,8	2,2	2,3–10,2
Társalgás	15,0	8,4	6,2–38,7
Idősek			
Narratíva	10,9	5,7	5,2–30,1
Visszamondás	8,4	4,3	3,1–18,1
Társalgás	15,3	5,4	7,1–32,9

Csak a nők beszédprodukcióit elemezve azt találtuk, hogy a megakadások gyakoriságát az életkor nem, míg a beszéd típus nagymértékben befolyásolta: $F(2, 58) = 20,943$; $p \leq 0,001$; $\eta^2 = 0,437$. Mind a fiataloknál, mind az idősek-nél a tartalomösszegzés és a társalgás között volt szignifikáns a különbség ($p \leq 0,001$).

Csak a férfiak beszédprodukcióit összevetve a statisztikai elemzés mind az életkor (az idősek beszéde folyamatosabb volt) [$F(1, 59) = 10,026$; $p = 0,003$; $\eta^2 = 0,157$], mind a beszéd típus [$F(2, 58) = 5,997$; $p = 0,004$; $\eta^2 = 0,182$] szignifikáns hatását igazolta. A beszédprodukciókat életkoronként is összevetve a post hoc teszt a fiatal férfiak tartalomösszegzései és társalgásai között mutatott matematikai eltérést ($p = 0,036$).

Összevetettük az összes megakadáson belül a bizonytalanságra visszavezethető jelenségek és a hibák arányát. Mindegyik beszéd típusban, életkorban és mindkét nemnél a megakadások döntő többsége (legalább 84%-a) bizonytalansági megakadás volt (2. ábra). Ugyanazon életkorban és beszéd típusban a két nem összes megakadásainak összevetésekor (az egyéni különbségeket nem figyelembe véve) csak az idősek tartalomösszegzésében figyeltünk meg nagyobb különbséget a nők és a férfiak között: az idős férfiaknál 8,3 százalékponttal több hiba típusú megakadás fordult elő.



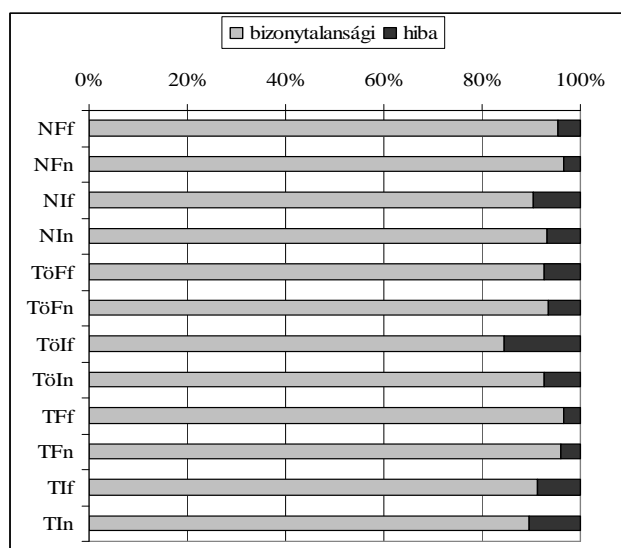
1. ábra

A beszédprodukciók folyamatossága (az összes megakadás gyakorisága) az életkor, a nem és a beszéd típus függvényében
(F = fiatal, I = idős; f = férfi, n = nő)

A bizonytalansági megakadások elemzésekor az összes megakadásnál adatszámokhoz hasonló eredményeket kaptunk: mindegyik beszéd típusban az idősek beszéde volt folyamatosabb, azaz a fiataloknál gyakoribbak voltak az elemzett jelenségek. Az egyéni különbségeket nem tekintetbe véve a csoportokra jellemző gyakorisági értékek a következők voltak: a narratívában a fiataloknál 8,2 szavanként, az időseknél 10,0 szavanként; a tartalomösszegzésben a fiataloknál 5,6 szavanként, az időseknél 7,4 szavanként; a társalgásban a fiataloknál 12,5 szavanként, az időseknél 15,8 szavanként fordult elő megakadásjelenség. Ugyanakkor az egyéni eredményeket kiszámítva azt találtuk, hogy a szórás a narratíva és a tartalomösszegzés esetén az időseknél több mint kétszeres volt, mint a fiataloknál (2. táblázat). A következő példában egy idős nő tartalomösszegzésében előforduló bizonytalansági megakadások (átlagosnál nagyobb) gyakoriságát figyelhetjük meg (a számok a néma szünetek időtartamát, a zárójelben levő számok a hezitálások időtartamát jelzik):

- (1) tessék nézni **ahogy** 50 **ahogy** **aa** 412 **a** 72 oroszlán 634 **önnő** (1231) 369 **öö** (476) jött **öö** (92) felé **öö** (225) nem tom **milye milyen** állat ami meg akarta támadni és az oroszlán **ugye** 416 **ahogy** **öö** (176) 512

öm (182) 237 fölfedezte 297 s ahogy fogta a kicsinyét és **má már** vitte tovább.



2. ábra

A kétféle megakadástípus aránya az összes megakadáshoz képest az életkor, a nem és a beszéd típus függvényében

(N = narratíva, Tö = tartalomösszegzés, T = társalgás; F = fiatal, I = idős; f = férfi, n = nő)

A statisztikai elemzés szerint a bizonytalansági megakadások gyakoriságára is a beszéd típus volt a legnagyobb hatással: $F(2, 118) = 14,697$; $p \leq 0,001$; $\eta^2 = 0,214$. Önmagában az életkor [$F(1, 119) = 6,805$; $p = 0,010$; $\eta^2 = 0,059$], az életkor és a nem együttesen [$F(1, 119) = 6,663$; $p = 0,011$; $\eta^2 = 0,058$], illetve a beszéd típus és a nem is [$F(2, 118) = 3,817$; $p = 0,025$; $\eta^2 = 0,066$] befolyásolta a gyakorisági értékeket.

A nők és a férfiak között csak az idős beszélők tartalomösszegzésében találtunk szignifikáns különbséget [$F(1, 19) = 7,333$; $p = 0,014$]: az idős férfiak ritkábban produkáltak bizonytalansági megakadást, mint az idős nők. A narratívákban a fiatal nőknél átlagosan 11,4 szavanként (a szórás: 4,8), a fiatal férfiaknál 7,7 szavanként (a szórás: 3,2); az idős nőknél átlagosan 10,6 szavanként (a szórás: 4,3), az idős férfiaknál 14,6 szavanként (a szórás: 12,1) fordult elő bizonytalansági megakadás. A tartalomösszegzésekben a fiatal nőknél átlagosan 6,2 szavanként (a szórás: 1,6), a fiatal férfiaknál 6,2 szavanként (a szórás: 3,1); az idős nőknél átlagosan 6,6 szavanként (a szórás:

3,1), az idős férfiaknál 15,1 szavanként (a szórás: 9,4) fordult elő bizonytalansági megakadás. A társalgásokban a fiatal nőknél átlagosan 18,2 szavanként (a szórás: 10,2), a fiatal férfiaknál 12,9 szavanként (a szórás: 6,1); az idős nőknél átlagosan 18,7 szavanként (a szórás: 7,5), az idős férfiaknál 16,0 szavanként (a szórás: 5,1) fordult elő bizonytalansági megakadás.

2. táblázat: A beszéd folyamatossága (a bizonytalansági megakadások gyakorisága) az életkor és a beszéd típus függvényében (szó/bizonytalansági megakadás; minél magasabb az érték, annál folyamatosabb a beszéd)

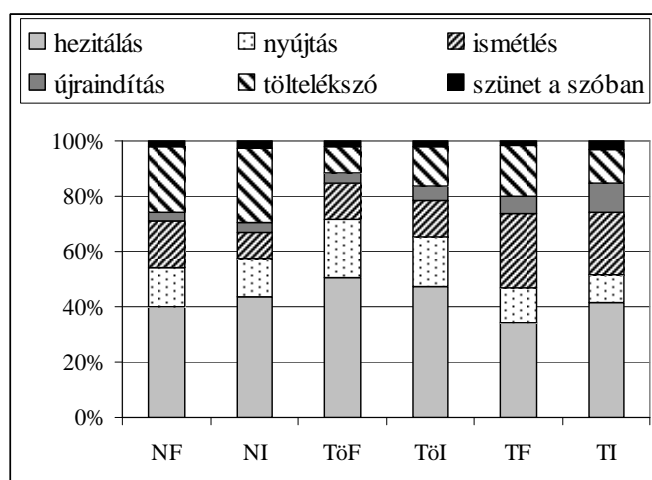
Beszéd típus	Átlag	Átlagos eltérés	Minimum–maximum
Fiatalok			
Narratíva	9,5	4,4	4,0–19,5
Visszamondás	6,2	2,4	2,4–10,9
Társalgás	15,6	8,6	6,2–38,7
Idősek			
Narratíva	12,6	9,1	5,3–47,7
Visszamondás	10,9	8,1	3,2–37,9
Társalgás	17,3	6,4	7,2–37,3

Csak a nők beszédprodukcióit összevetve azt találtuk, hogy a beszéd típus volt szignifikáns hatással a gyakorisági értékekre [$F(2, 58) = 20,658; p \leq 0,001; \eta^2 = 0,433$], az életkor nem. A post hoc teszt a fiatal nőknél a tartalomösszegzés és a társalgás között igazolt szignifikáns különbséget ($p = 0,001$), az idős nőknél pedig a társalgás és a másik két beszéd típus között ($p \leq 0,039$).

Csak a férfiak beszédprodukcióit elemezve az életkor esetében igazolódott szignifikáns hatás: $F(1, 59) = 11,256; p = 0,001; \eta^2 = 0,172$. Az idős férfiak beszéde folyamatosabb volt, mint a fiatal férfiaké mindhárom beszéd típusban.

A bizonytalansági megakadások típusainak elemzése azt mutatta, hogy az életkortól és a beszéd típustól függetlenül a hezitálások fordulnak elő a legnagyobb arányban mindegyik beszéd típusban (3. ábra). Ennek egyrészt egyfajta nyelvhasználati „divat” lehet az oka, másrészt ezt a megakadástípust a leg-egyszerűbb előhívni a tervezési bizonytalanság esetén. Ugyanakkor vannak életkor-specifikus és az adott beszédhelyzetre jellemző(bb) megakadás-típusok. A fiataloknál például mindhárom beszéd típusban nagyobb arányban fordultak elő nyújtások, illetve két beszéd típusban ismétlések, mint az időseknél. Az idősekre azonban nagyobb arányban voltak jellemzőek mindegyik beszéd típusban az újraindítások (ezek a szóelőhívás bizonytalanságát jelzik, vö. Gyarmathy 2009), illetve a monologikus szövegekben a töltelékszók előfordulása. Szintén jellemzőbb volt az időseknél a narratívákban és a társalgásban a szünet a szóban jelenség, amely a tartalomösszegzésben azonos arányban jelent meg mindkét életkori csoportban. A nemek közötti összeve-

tés szerint az életkortól és a beszéd típustól is függött, hogy melyik megakadási típus melyik nemnél jelent meg nagyobb arányban (4. ábra).

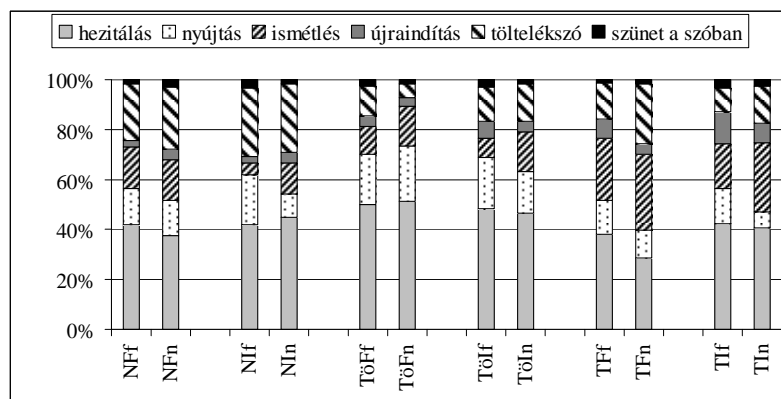


3. ábra

A bizonytalansági megakadások típusainak aránya
(100% az adott beszéd típusban és életkori csoportban adatolt összes bizonytalansági megakadás; N = narratíva, Tö = tartalomösszegzés, T = társalgás; F = fiatal, I = idős)

Ezek a százalékos arányok nem mutatják meg azt, hogy a szövegben az összes szószámhoz viszonyítva milyen gyakori volt egy-egy megakadási típus, viszont arra következtethetünk belőlük, hogy ha az adott beszélői csoport egy bizonyos beszéd típusban bizonytalansági megakadást produkál, akkor milyen arányban jelenik meg az adott megakadási típus.

Az egyes bizonytalansági megakadási típusok szószámra vetített gyakorisága nagy különbségeket mutatott mind a beszéd típusok, mind az életkorok és a nemek szerint is. Bár a 3. és a 4. ábrán feltüntetett arányok azt mutatják, hogy vannak inkább a fiatalokra, illetve inkább az idősekre jellemző típusok, az egyéni különbségeket nem tekintve a csoportokra jellemző gyakorisági értékek azt mutatják, hogy mindhárom beszéd típusban mindegyik típus a fiataloknál fordul elő gyakrabban, ez alól mindössze négy esetben volt kivétel. Az időseknél gyakoribb volt a tartalomösszegzéskor a töltelékszó és az újraindítás, mint a fiataloknál, illetve a társalgásban a szünet a szóban jelenség és az újraindítás. (A konkrét értékeket a statisztikai elemzésnél közöljük.)



4. ábra

A bizonytalansági megakadások típusainak aránya a két nemnél
(N = narratíva, Tő = tartalomösszegzés, T = társalgás; F = fiatal, I = idős;
f = férfi, n = nő)

A megakadástípusok többsége nem fordult elő minden adatközlőnél minden beszéd típusban. A 3. táblázatban feltüntetett adatok azt mutatják, hogy azoknál a beszélőknél, akiknél előfordult az adott megakadástípus, milyen gyakorisággal, hány szavanként jelent meg (azaz ezekben az értékekben nem jelennek meg azok az adatközlők, akiknél nem adatoltuk az adott megakadástípust). Az egyetlen mindenkinél, mindegyik beszédprodukcióban megjelenő megakadás a hezitálás volt; míg az újraindítás és a szünet a szóban jelenség gyakorisági mutatóin nem tudtunk statisztikai elemzést végezni, mert ezek a típusok sok beszédprodukcióban nem jelentek meg.

Hezitálást csoportszinten a narratívákban a fiataloknál 20,5 szavanként, az időseknél 22,7 szavanként; a tartalomösszegzésekben a fiataloknál 11,0 szavanként, az időseknél 15,7 szavanként; a társalgásban a fiataloknál 36,5 szavanként, az időseknél 38,2 szavanként adatoltunk. Az egyéni eredményeket összevető statisztikai elemzés szerint a hezitálás gyakoriságát a beszéd típus $[F(2, 118) = 14,957; p \leq 0,001; \eta^2 = 0,218]$, illetve az életkor és a nem együttesen $[F(1, 119) = 5,012; p = 0,027; \eta^2 = 0,045]$ befolyásolta szignifikánsan. A post hoc teszt szerint a társalgásokban előforduló hezitálások gyakorisága különbözött szignifikánsan a másik két beszéd típusban adatolt gyakoriságtól ($p \leq 0,001$).

Ismétlést csoportszinten a narratívákban a fiataloknál 50,0 szavanként, az időseknél 110,9 szavanként; a tartalomösszegzésekben a fiataloknál 42,8 szavanként, az időseknél 55,8 szavanként; a társalgásban a fiataloknál 46,3 szavanként, az időseknél 70,1 szavanként adatoltunk. Az ismétlések gyakoriságára a beszéd típus és az életkor együttesen $[F(2, 106) = 3,237; p = 0,044; \eta^2$

= 0,063], illetve az életkor és a nem együttesen [$F(1, 107) = 7,974$; $p = 0,006$; $\eta^2 = 0,077$] volt szignifikáns hatással.

3. táblázat: A bizonytalansági megakadások típusainak gyakorisága (átlag és átlagos eltérés) (szó/megakadás; minél magasabb az átlag, annál ritkább az adott megakadástípus; a táblázatban csak azon beszélők adatai szerepelnek, akiknél előfordultak az adott megakadástípusok)

Beszéd típus	Fiatal férfi	Fiatal nő	Idős férfi	Idős nő
Hezitálás				
Narratíva	19,7; 11,2	32,8; 15,9	38,5; 27,4	27,4; 20,3
Visszamondás	12,8; 6,8	15,4; 15,0	44,1; 36,5	16,8; 13,3
Társalgás	40,5; 27,9	64,4; 39,4	52,3; 33,5	62,2; 50,0
Nyújtás				
Narratíva	93,5; 116,7	111,9; 65,2	76,6; 42,2	111,6; 52,5
Visszamondás	61,3; 82,0	38,8; 24,7	109,0; 91,0	44,9; 23,8
Társalgás	150,8; 189,9	221,7; 202,4	198,0; 164,2	304,2; 170,1
Ismétlés				
Narratíva	54,0; 29,9	112,0; 126,1	282,9; 144,7	100,1; 64,9
Visszamondás	131,5; 153,5	88,6; 105,3	137,5; 110,9	58,7; 50,5
Társalgás	89,5; 109,6	189,1; 282,6	131,0; 115,6	75,5; 37,9
Újraindítás				
Narratíva	369,2; 229,6	358,5; 231,0	384,2; 181,8	330,1; 231,8
Visszamondás	114,9; 86,4	146,8; 90,6	183,9; 148,2	99,8; 26,6
Társalgás	152,2; 52,1	409,8; 256,0	134,0; 61,0	234,0; 140,2
Töltelékszó				
Narratíva	29,6; 6,7	81,4; 68,9	56,5; 48,8	57,7; 55,7
Visszamondás	61,2; 32,1	109,2; 77,8	150,4; 127,2	78,3; 76,9
Társalgás	160,5; 191,0	149,9; 141,6	200,0; 140,8	183,1; 147,7
Szünet a szóban				
Narratíva	450,0; 261,9	377,3; 304,6	289,5; 176,2	463,9; 206,1
Visszamondás	250,9; 142,7	263,7; 169,0	221,7; 106,7	226,7; 224,3
Társalgás	543,2; 181,9	412,7; 163,0	344,0; 118,0	612,1; 210,0

Töltelékszót csoportszinten a narratívákban a fiataloknál 34,7 szavanként, az időseknél 36,6 szavanként; a tartalomösszegzésekben a fiataloknál 58,6 szavanként, az időseknél 51,1 szavanként; a társalgásban a fiataloknál 68,1 szavanként, az időseknél 129,8 szavanként adatoltunk. Nyújtást csoportszinten a narratívákban a fiataloknál 57,6 szavanként, az időseknél 71,6 szavanként; a tartalomösszegzésekben a fiataloknál 26,2 szavanként, az időseknél 41,0 szavanként; a társalgásban a fiataloknál 98,9 szavanként, az időseknél 152,3 szavanként adatoltunk. A beszéd típus szignifikáns hatása igazolódott mind a töltelékszavak gyakoriságában [$F(2, 112) = 11,685$; $p \leq 0,001$; $\eta^2 = 0,186$], mind a nyújtások [$F(2, 115) = 18,065$; $p \leq 0,001$; $\eta^2 = 0,256$] előfordulásában. A post hoc teszt szerint mindkét esetben a társalgásnál mért

gyakorisági adatok különböztek szignifikánsan a másik két beszédtypustól a töltelékszavak esetén $p \leq 0,008$, a nyújtások esetén $p \leq 0,001$.

Újraindítás csoportszinten a narratívákban a fiataloknál 258,4 szavanként, az időseknél 264,1 szavanként; a tartalomösszegzésekben a fiataloknál 147,2 szavanként, az időseknél 141,7 szavanként; a társalgásban a fiataloknál 201,0 szavanként, az időseknél 150,3 szavanként fordult elő. Szünet a szóban jelenség csoportszinten a narratívákban a fiataloknál 391,7 szavanként, az időseknél 410,9 szavanként; a tartalomösszegzésekben a fiataloknál 273,4 szavanként, az időseknél 358,0 szavanként; a társalgásban a fiataloknál 830,7 szavanként, az időseknél 544,0 szavanként jelent meg.

A hibák életkorra jellemző gyakoriságát tekintve ellentétes eredményeket kaptunk a bizonytalansági megakadásokhoz képest: mindegyik beszédtypusban az idősök produkáltak gyakrabban hibát, és náluk voltak kisebbek az egyének közötti különbségek is (4. táblázat). A csoportátlagok a következőképpen alakultak: a narratívában a fiataloknál 189,8 szavanként, az időseknél 115,6 szavanként; a tartalomösszegzésben a fiataloknál 74,6 szavanként, az időseknél 66,7 szavanként; a társalgásban a fiataloknál 336,8 szavanként, az időseknél 150,3 szavanként adatoltunk hibát.

4. táblázat: A hibák gyakorisága az életkor és a beszédtypus függvényében (szó/hiba, minél magasabb az érték, annál ritkább a hiba; a táblázatban csak azon beszélők adatai szerepelnek, akiknél előfordultak hibák)

Beszédtypus	Átlag	Átlagos eltérés	Minimum–maximum
Fiatalok			
Narratíva	232,2	144,0	78,9–590,0
Visszamondás	106,6	70,3	12,2–318,0
Társalgás	348,6	167,7	126,0–673,0
Idősök			
Narratíva	167,0	131,1	38,7–644,0
Visszamondás	102,4	66,4	13,3–241,0
Társalgás	203,5	123,0	67,1–486,0

Mindhárom beszédtypusban volt olyan fiatal beszélő, akinél egyetlen hiba típusú megakadás sem fordult elő. Az egyéni gyakorisági mutatók szerint a hibák előfordulása is nagymértékben függött a beszédtypustól, a leggyakrabban a tartalomösszegzésekben, a legritkábban a társalgásban fordultak elő. A statisztikai elemzés szerint az életkor [$F(1, 109) = 9,906$; $p = 0,002$; $\eta^2 = 0,092$] és a beszédtypus [$F(2, 108) = 17,519$; $p \leq 0,001$; $\eta^2 = 0,263$] hatása is szignifikáns volt a gyakoriságra. A post hoc teszt szerint mindegyik beszédtypus szignifikánsan különbözött a másik kettőtől ($p \leq 0,040$).

A nők és a férfiak között egyik beszédtypusban és életkori csoportban sem találtunk szignifikáns különbséget; a következő átlagok kiszámításakor, illetve a statisztikai elemzéskor csak azoknak az adatközlőknek az adatait vettük

figyelembe, akiknél előfordult hiba. A narratívákban a fiatal nőknél átlagosan 292,7 szavanként (a szórás: 173,6), a fiatal férfiaknál 183,8 szavanként (a szórás: 99,4); az idős nőknél átlagosan 191,0 szavanként (a szórás: 171,4), az idős férfiaknál 142,9 szavanként (a szórás: 75,0) fordult elő hiba típusú megakadás. A tartalomösszegzésekben a fiatal nőknél átlagosan 80,5 szavanként (a szórás: 25,9), a fiatal férfiaknál 135,9 szavanként (a szórás: 93,1); az idős nőknél átlagosan 102,7 szavanként (a szórás: 69,2), az idős férfiaknál 102,1 szavanként (a szórás: 67,2) fordult elő hiba. A társalgásokban a fiatal nőknél átlagosan 348,0 szavanként (a szórás: 147,2), a fiatal férfiaknál 349,1 szavanként (a szórás: 196,4); az idős nőknél átlagosan 207,8 szavanként (a szórás: 137,3), az idős férfiaknál 200,5 szavanként (a szórás: 114,3) fordult elő bizonytalansági megakadás.

Csak a nők beszédprodukcióit összevetve azt találtuk, hogy mind az életkor [$F(1, 54) = 4,235$; $p = 0,045$; $\eta^2 = 0,080$], mind a beszéd típus szignifikáns hatással volt a gyakorisági értékekre [$F(2, 53) = 10,480$; $p \leq 0,001$; $\eta^2 = 0,300$], ez utóbbi hatása volt a nagyobb. A post hoc teszt a fiatal nőknél igazolt a tartalomösszegzés és másik két beszéd típus között szignifikáns különbséget ($p \leq 0,020$).

Csak a férfiak beszédprodukcióinak elemzésekor (a nőkhez hasonlóan) a beszéd típus nagyobb mértékű hatása igazolódott [$F(2, 53) = 8,891$; $p = 0,001$; $\eta^2 = 0,266$], de az életkor is szignifikáns hatással volt a gyakorisági értékekre [$F(1, 54) = 5,926$; $p = 0,019$; $\eta^2 = 0,108$]. A post hoc teszt a fiatal férfiaknál igazolt a társalgás és a másik két beszéd típus között szignifikáns különbséget ($p \leq 0,037$).

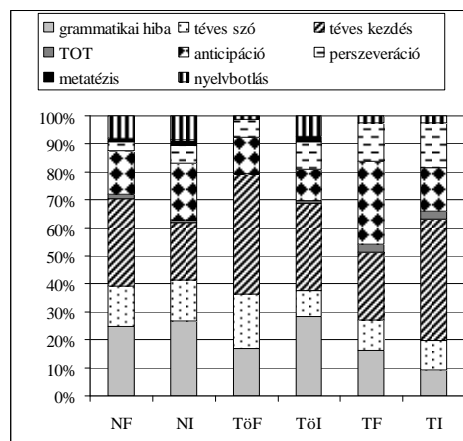
A hibatípusok elemzése azt mutatta, hogy életkortól és beszéd típustól függetlenül a szóelőhívás szintjén fordul elő a legtöbb probléma, illetve vannak inkább az idősekre jellemző hibák (5. ábra). A perszeveráció például mindegyik beszéd típusban az időseknél jelent meg nagyobb arányban. A két nem között is megfigyelhetők különbségek a megakadások arányában, de ezt az életkor és a beszéd típus is befolyásolta (6. ábra). Az ábrákon nem tüntettük fel a kontaminációt, amely mindössze kétszer fordult elő az összes elemzett beszéd produkcióban.

Az egyes hibatípusok szószámra vetített gyakorisága nagy különbségeket mutatott mind a beszéd típusok, mind az életkorok és a nemek szerint is (5. táblázat; a táblázatban csak azoknak a beszélőknek az adatait tüntettük fel, akiknél előfordult az adott tervezési szinthez köthető hibatípus). Egyetlen olyan hibatípus sem volt, amely minden adatközlőnél minden beszéd típusban előfordult volna. Mivel az egyes hibatípusok megjelenése igen ritka volt, ezért a beszéd tervezési szinteknek megfelelő hibák együttesének gyakoriságát vizsgáltuk meg.

A nyelvi tervezéshez köthető hibát (grammatikai hiba, kontamináció) csoportszinten a narratívákban a fiataloknál 759,0 szavanként, az időseknél 410,9 szavanként; a tartalomösszegzésekben a fiataloknál 441,6 szavanként, az időseknél 226,7 szavanként; a társalgásban a fiataloknál 2076,8 szavan-

ként, az időseknél 1632,0 szavanként adatoltunk. Az egyéni különbségeket is figyelembe vevő statisztikai elemzés szerint (amelyben csak azoknak a beszélőknek az adatait vettük figyelembe, akiknél megjelent az adott hibatípus) a nyelvi tervezéshez köthető hibák (grammatikai hiba és kontamináció) gyakoriságát csak a beszéd típus befolyásolta szignifikánsan [$F(2, 53) = 15,427$; $p \leq 0,001$; $\eta^2 = 0,418$]. A post hoc teszt szerint a tartalomösszegzésben előforduló (nyelvi tervezéshez köthető) hibák gyakorisága különbözött szignifikánsan a másik két beszéd típusban adatolt gyakoriságtól ($p \leq 0,001$). A következő példában egy grammatikai hiba figyelhető meg (az idős férfi beszélő a téves toldalékot javítja):

(2) városnak a védelmében 559 a **polgárokat** 136 **nak** is fegyvert adott



5. ábra

A hibák típusainak aránya

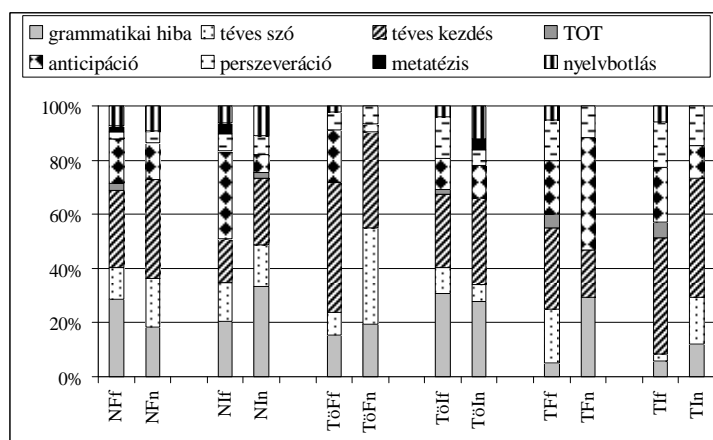
(100% az adott beszéd típusban és életkori csoportban adatolt összes hiba; N = narratíva, Tö = tartalomösszegzés, T = társalgás; F = fiatal, I = idős)

A lexikális hozzáférés problémáit (téves szó, téves kezdés és „nyelvem hegyén van” jelenség) csoportszinten a narratívákban a fiataloknál 809,6 szavanként, az időseknél 326,3 szavanként; a tartalomösszegzésekben a fiataloknál 119,6 szavanként, az időseknél 174,4 szavanként; a társalgásban a fiataloknál 890,1 szavanként, az időseknél 265,7 szavanként adatoltuk. A szóelőhívás problémáinak gyakoriságára hatással volt a beszéd típus [$F(2, 88) = 32,292$; $p \leq 0,001$; $\eta^2 = 0,453$] és az életkor is [$F(1, 89) = 14,183$; $p \leq 0,001$; $\eta^2 = 0,154$] külön-külön, illetve együttesen is [$F(2, 88) = 7,289$; $p = 0,001$; $\eta^2 = 0,157$], a post hoc teszt szerint a beszéd típust tekintve a tartalomösszeg-

zésben előforduló szóelőhívási hibák gyakorisága különbözött szignifikánsan a másik két beszéd típusban adatolt gyakoriságtól ($p \leq 0,001$).

Egy idős nő tartalomösszegzéséből származik a következő téves kezdés (egy idős nő tartalomösszegzéséből az *é* hang a téves kezdés):

(3) megeskette 151 öket 551 arra hogy 396 *é* hogy 411 halálukig



6. ábra

A hibák típusainak aránya a két nemnél

(N = narratíva, Tö = tartalomösszegzés, T = társalgás; F = fiatal, I = idős; f = férfi, n = nő)

Sorrendiségi hibát (anticipáció, perszeveráció, metatézis) csoportszinten a narratívákban a fiataloknál 934,2 szavanként, az időseknél 410,9 szavanként; a tartalomösszegzésekben a fiataloknál 382,7 szavanként, az időseknél 272,1 szavanként; a társalgásban a fiataloknál 778,8 szavanként, az időseknél 476,0 szavanként adatoltunk. A beszéd típus szignifikáns hatása igazolódott a sorrendiségi hibák gyakoriságára is [$F(2, 60) = 5,424$; $p = 0,007$; $\eta^2 = 0,178$], a post hoc teszt szerint a tartalomösszegzésben előforduló sorrendiségi hibák gyakorisága különbözött szignifikánsan a másik két beszéd típusban adatolt gyakoriságtól ($p \leq 0,023$). A következő példában az idős adatközlő előbb ejti ki a *visel* szót, mint annak a közlésben szerepelnie kellene (anticipáció):

(4) viselték a 335 vagy öö (165) figyelték a viselkedésüket

Az artikulációs kivitelezés hibáját (egyszerű nyelvbtlás) csoportszinten a narratívákban a fiataloknál 2428,8 szavanként, az időseknél 1386,8 szavanként; a tartalomösszegzésekben a fiataloknál 5741,0 szavanként, az időseknél 850,3 szavanként; a társalgásban a fiataloknál 12461,0 szavanként, az időseknél 5712,0 szavanként adatoltuk. Ennek a hibatípusnak a gyakoriságára a

kevés adat miatt (a társalgásban például mindössze két idős és egy fiatal férfinél fordult elő nyelvbottlás; a fiatalok tartalomösszegzésében összesen egy főnél jelent meg ez a megakadás) nem tudtuk elvégezni a statisztika próbát.

5. táblázat: Az egyes hibatípusok gyakorisága (átlag és átlagos eltérés) (szó/hiba; minél magasabb az átlag, annál ritkább az adott hiba; a táblázatban csak azon beszélők adatai szerepelnek, akiknél előfordultak az adott hibatípusok)

Beszéd típus	Fiatal férfi	Fiatal nő	Idős férfi	Idős nő
A nyelvi tervezés hibái				
Narratíva	366,3; 154,8	565,3; 126,8	455,6; 151,8	381,8; 201,1
Visszamondás	370,7; 268,8	174,0; 53,9	173,0; 112,6	228,2; 79,0
Társalgás	606,0; –	528,3; 121,6	564,5; 78,5	638,3; 206,4
A szóelőhívás hibái				
Narratíva	450,6; 267,6	607,7; 110,6	364,1; 152,8	343,2; 156,0
Visszamondás	169,1; 92,6	114,0; 47,2	219,8; 145,0	165,2; 94,0
Társalgás	519,1; 218,5	540,0; 31,3	332,3; 152,4	303,4; 195,4
Sorrendiségi hibák				
narr.	538,0; 216,0	498,0; 145,7	233,4; 103,9	533,2; 137,7
visszam.	258,4; 136,0	312,3; 143,4	221,5; 109,0	272,0; 265,3
társ.	467,8; 240,0	445,7; 199,2	390,8; 184,7	414,9; 250,9
az artikulációs kivitelezés hibája				
Narratíva	522,3; 377,2	617,0; 60,8	598,7; 11,4	512,9; 212,2
Visszamondás	546,0; –	–	349,5; 119,5	281,2; 75,1
Társalgás	799,0; –	–	616,5; 154,9	–

Összegzés, következtetések

Tanulmányunkban a megakadásjelenségek gyakoriságát és típusait elemeztük a beszéd típus, az életkor és a nem függvényében. A beszéd típus hatására vonatkozó első hipotézisünk igazolódott (a vizsgált tényezők közül ennek a hatása volt a legnagyobb a gyakorisági értékekre): a narratívákhoz képest mindkét életkori csoport beszédében több megakadás jelent meg a nagyobb mentális erőfeszítést igénylő tartalomösszegzésekben, illetve csökkent a megakadások gyakorisága a rövidebb beszédlépésekből álló, a tervezésre több időt biztosító társalgásokban.

Bár kisebb mértékben, de az életkor is hatással volt a megakadások előfordulására. Ugyanakkor a második hipotézisünk csak részben igazolódott, mert a várttal ellentétben az időseknél ritkábban fordultak elő megakadásjelenségek, mint a fiataloknál, mindhárom beszéd típusban. A megakadástípusok elemzése ugyanakkor igazolta azt, hogy vannak inkább az idősekre, illetve inkább a fiatalokra jellemző megakadások: amíg a fiataloknál gyakoribbak

voltak a bizonytalanságra visszavezethető jelenségek, addig az időseknél szignifikánsan gyakoribb volt a hibák megjelenése.

A három tényező közül a legkisebb (elenyésző) hatása a nemnek volt a gyakoriságra. Tendenciaszerűen a fiataloknál a nők, az időseknél a férfiak produkáltak ritkábban megakadást, de a gyakorisági értékeket erőteljesen befolyásolta a beszéd típus és a megakadástípus is.

Az eredmények arra hívják fel a figyelmet, hogy a megakadások gyakoriságának elemzésekor mindig figyelembe kell venni a beszéd típust és az életkort. Az eltérő mentális erőfeszítést igénylő beszédhelyzetek mind a fiatalok, mind az idősek esetében különböző beszédprodukciós stratégiákat igényelnek. Ugyanakkor az életkor önmagában is hat a beszédre: a lassabb működések, a nehezített szóelőhívás és nyelvi tervezés befolyásolja az idősek beszédében megjelenő hibákat; míg a fiatalok gyorsabb tempója, s ezáltal a gondolkodásra szánt kevesebb idő a bizonytalansági megakadások gyakoribb előfordulásában mutatkozik meg. A kutatás során tapasztalt életkori különbségek tehát nem azt jelentik, hogy az idősek beszédtervezési folyamatai gyengébbek lennének a fiatalokénál, hanem inkább egy másféle működésre utalnak.

Irodalom

- Bakti Mária 2009. Megakadások a szinkrontolmácsok spontán beszédében. *Beszédkutatás* 2009. 217–227.
- Bóna Judit 2006. Tudunk-e változtatni spontán beszédünk tempóján? In Mártonfi Attila – Papp Kornélia – Slíz Mariann (szerk.): *101 írás Pusztai Ferenc tiszteletére*. Argumentum Kiadó, Budapest. 560–566.
- Bóna Judit 2007. Magánhangzó-nyújtások akusztikai-fonetikai paraméterei a spontán beszédben. *Beszédkutatás* 2007. 99–107.
- Bóna Judit 2012. A spontán beszéd sajátosságai idősödő, idős és matuzsálemi korban. In Markó Alexandra (szerk.): *Beszédtudomány. Az anyanyelv-elsajátítástól a zöngelkedési időig*. ELTE BTK–MTA Nyelvtudományi Intézet, Budapest. 100–115.
- Bóna Judit 2013a. A beszéd szünetek fonetikai sajátosságai a beszéd típus függvényében. *Beszédkutatás* 2013. 60–75.
- Bóna Judit 2013b. Narrative recall in the elderly: Content, fluency and speech errors in the narrative speech of young, young-old and old-old speakers. *Acta Linguistica Hungarica* 60/2. 123–142.
- Bortfeld, Heather – Leon, Silvia D. – Bloom, Jonathan E. – Schober, Michael F. – Brennan, Susan E. 2001. Disfluency rates in conversation: effects of age, relationship, topic, role, and gender. *Language and Speech* 44/2. 123–147.
- Branigan, Holly – Lickley, Robin – McKelvie, David 1999. Non-linguistic influences on rates of disfluency in spontaneous speech. In *Proceedings of the 14th International Conference of Phonetic Sciences*. University of California. 387–390.
- Burke, Deborah M. – MacKay, Donald G. – Worthley, Joanna S. – Wade, Elizabeth 1991. On the tip of the tongue: What causes word finding failures in young and older adults. *Journal of Memory and Language* 30. 542–579.

- Duchin, Sandra W. – Mysak, Edward D. 1987. Disfluency and rate characteristics of young adult, middle-aged, and older males. *Journal of Communication Disorders* 20. 245–257.
- Fromkin, Victoria A. 1973. The non-anomalous nature of anomalous utterances. In Victoria A. Fromkin (ed.): *Speech errors as linguistic evidence*. Mouton, The Hague. 215–242.
- Gocsál Ákos 2001. Gyorsabban beszélnek-e a nők, mint a férfiak? *Beszéd kutatás 2001*. 61–72.
- Goldman-Eisler, Frieda 1958. Speech analysis and mental processes. *Language and Speech* 1. 59–75.
- Gósy Mária 2000. A beszéd szünetek kettős funkciója. *Beszéd kutatás 2000*. 1–15.
- Gósy Mária 2003. A spontán beszédben előforduló megakadásjelenségek gyakorisága és összefüggései. *Magyar Nyelvőr* 127. 257–277.
- Gósy Mária 2005. *Pszicholingvisztika*. Osiris Kiadó, Budapest.
- Gósy Mária 2010. Szövegértés alapú narratívák. In Bárdosi Vilmos (szerk.): *Világkép a nyelvben és a nyelvhasználatban*. Tinta Kiadó, Budapest. 113–124.
- Gósy Mária – Bóna Judit – Beke András – Horváth Viktória 2013. A kitöltött szünetek fonetikai sajátosságai az életkor függvényében. *Beszéd kutatás 2013*. 121–143.
- Gósy Mária – Bóna Judit – Grácz Tekla Etelka – Gyarmathy Dorottya – Horváth Viktória – Imre Angéla – Markó Alexandra – Neuberger Tilda (szerk.) 2009. „Nyelvbotlás”-korpusz. 6. rész. *Beszéd kutatás 2009*. 257–267.
- Gósy Mária – Gyarmathy Dorottya – Horváth Viktória – Grácz Tekla Etelka – Beke András – Neuberger Tilda – Nikléczy Péter 2012. BEA: Beszélt nyelvi adatbázis. In Gósy Mária (szerk.): *Beszéd, adatbázis, kutatások*. Akadémiai Kiadó, Budapest. 9–24.
- Gósy Mária – Horváth Viktória 2009. Hogyan tükrözi a kiejtés a nyelvi funkció változását? In Keszler Borbála – Tátrai Szilárd (szerk.): *Diskurzus a grammatikában – grammatika a diskurzusban*. Tinta Könyvkiadó, Budapest. 37–45.
- Gyarmathy Dorottya 2009. A beszélő bizonytalanságának jelzései: ismétlések és újra-indítások. *Beszéd kutatás 2009*. 196–216.
- Hnath-Chisolm, Theresa – Willott, James F. – Lister, Jennifer J. 2003. The aging auditory system: Anatomic and physiologic changes and implications for rehabilitation. *International Journal of Audiology* 42. Supplement 2. 3–10.
- Horváth Viktória 2004. Megakadásjelenségek a párbeszédekben. *Beszéd kutatás 2004*. 187–199.
- Horváth Viktória 2005. The tip of the tongue phenomenon with elderly. In Lengyel, Zsolt – Navracsics, Judit (eds.): *Selected papers of 8th Summer School of Psycholinguistics*. Veszprémi Egyetem Alkalmazott Nyelvészeti Tanszéke, Veszprém.
- Horváth Viktória 2007. Vannak-e „női” és „férfi” megakadásjelenségek a spontán beszédben? *Magyar Nyelvőr* 131. 315–323.
- Horváth, Viktória 2010. Filled pauses in Hungarian: Their phonetic form and function. *Acta Linguistica Hungarica* 57/2–3. 288–306.
- Humes, Larry E. 1996. Speech understanding in the elderly. *Journal of the American Academy of Audiology* 7. 161–167.
- Huszár Ágnes 2009. *Bevezetés a gendernyelvészetbe*. Segédkönyvek a nyelvészet tanulmányozásához 98. Tinta Könyvkiadó, Budapest.

- Huszár Ágnes 2011. A magyarországi gendernyelvészeti kutatások története és perspektívái. *Alkalmazott Nyelvészeti Közlemények* 6/1. 73–80.
- Juncos-Rabadán, Onésimo – Pereiro, Arturo X. – Rodríguez, Maria Soledad 2005. Narrative speech in aging: Quantity, information content, and cohesion. *Brain and Language* 95. 423–434.
- Kemper, Susan 1992. Adults' sentence fragments: Who, what, when, where, and why. *Communication Research* 19. 445–458.
- Kemper, Susan – Rash, Shannon – Kynette, Donna – Norman, Suzanne 1990. Telling stories: The structure of adults' narratives. *European Journal of Cognitive Psychology* 2/3. 205–228.
- Laczkó Mária 2010. Megakadásjelenségek a spontán és a szónoki beszédben. *Beszéd-kutatás 2010*. 184–198.
- Leeper, Linda H. – Culatta, Richard 1995. Speech fluency: Effect of age, gender and context. *Folia Phoniatrica et Logopaedica* 47. 1–14.
- Levelt, Willem J. M. 1989. *Speaking. From intention to articulation*. A Bradford Book. The MIT Press, Cambridge (Massachusetts)–London (England).
- Markó Alexandra 2004. Megakadások vizsgálata különféle monologikus szövegekben. *Beszéd-kutatás 2004*. 209–222.
- Markó Alexandra 2005. A temporális szerkezet jellegzetességei eltérő kommunikációs helyzetekben. *Beszéd-kutatás 2005*. 63–77.
- Pethő Mária 2013. Megakadásjelenségek az extrovertáltak és introvertáltak prezentációiban. In Geecső Tamás – Sárdi Csilla (szerk.): *Az interkulturális kommunikáció elmélete és gyakorlata*. Segédkönyvek a nyelvészet tanulmányozásához 154. Kodolányi János Főiskola–Tinta Könyvkiadó, Székesfehérvár–Budapest, 178–183.
- Ramig, Lorraine Olson – Gray, Steven – Baker, Kristin – Corbin-Lewis, Kim – Buder, Eugene – Luschei, Erich – Coon, Hillary – Smith, Marshall 2001. The aging voice: A review, treatment data and familial and genetic perspectives. *Folia Phoniatrica et Logopaedica* 53. 252–265.
- Shriberg, Elizabeth 2001. To 'errrr' is human: Ecology and acoustics of speech disfluencies. *Journal of the International Phonetic Association* 31/1. 153–169.
- Szabó Eszter 2004. Megakadásjelenségek nyelvi játék közben. *Beszéd-kutatás 2004*. 200–208.
- Ulatowska, Hanna K. – Hayashi, Mari M. – Cannito, Michael P. – Flemig, Susan G. 1986. Disruption of reference in aging. *Brain and Language* 28. 24–41.
- Xue, Steve An – Hao, Grace Jianping 2003. Changes in the human vocal tract due to aging and the acoustic correlates of speech production: A pilot study. *Journal of Speech, Language, and Hearing Research* 46. 689–701.
- Yairi, Ehud – Clifton, Noel F. 1972. Disfluent speech behavior of preschool children, high school seniors, and geriatric persons. *Journal of Speech and Hearing Research* 15. 714–719.

AZ ARTIKULÁCIÓS TEMPÓ VIZSGÁLATA RÉGI MAGYAR FILMEKBEN

Gocsál Ákos

Bevezetés

A beszéd tempóviszonyaival foglalkozó szerzők gyakran említik, hogy beszédünk a 20. század folyamán felgyorsult. E megállapítások a leggyakrabban Hegedűs Lajos 1957-ben, illetve a Fónagy–Magdics szerzőpáros 1960-ban publikált adataira támaszkodnak, melyek szerint 1957-ben a hírolvasás tempója 12 hang/s, a sportközvetítésé pedig 14 hang/s volt, 1960-ban pedig több, különböző műfajú szövegből mérve átlagosan 11,35 hang/s adódott (az adatokat idézi: Kassai 1993). Mai méréseink során sokszor kapunk ezeknél magasabb tempóértékeket, emellett több olyan akusztikai fonetikai kísérlet is ismeretes, amelynek eredményeiből szintén a beszéd gyorsulására lehet következtetni. Ma egy perc alatt átlagosan húsz szóval ejtünk kitöbbet, mint négy évtizeddel ezelőtt (Gósy 2004), illetve a *tát*, *tátog*, *tátogat*, *tátogatók* szavaknál mért egyre rövidebb időtartamokból is a gyorsabb ejtésre lehet következtetni (Kassai 1993).

Újabb kutatási eredmények fényében azonban úgy tűnik, érdemes újból, más szempontok alapján is megvizsgálni a beszédtempó gyorsulásának problémakörét. Egyrészt azért, mert az utóbbi húsz évben lényegesen gyarapodott tudásunk azzal kapcsolatban, hogy egyéni sajátosságok, tulajdonságok hogyan differenciálják a beszéd akusztikai szerkezetét (pl. Horváth 2007, Bóna 2010, Gósy és mtsai 2013). Mivel a beszédakusztikai mérések igen aprólékos mérési munkát igényelnek, a leggyakrabban ma is csak néhány beszélőn végzett esettanulmányokat közölnek a szerzők, általános tendenciákra pedig számos ilyen esettanulmány együttes elemzésével következtetünk. Mindezek figyelembevételével újragondolhatjuk Kassai (1993) következtetéseit. Ugyanis ha egy konkrét személy 1909-ben viszonylag lassan, egy másik személy 1965-ben gyorsabban, egy harmadik, közelebbről meg nem határozott tulajdonságokkal rendelkező személy 1993-ban pedig még gyorsabban ejtette ki a *tát*, *tátog*, *tátogat*, *tátogatók* hangsorokat, ez önmagában még nem igazolja, hogy beszédünk tempója a teljes nyelvközösség vonatkozásában gyorsult volna a 20. század elejétől kezdődően. Ilyen megállapítást csak akkor tehetnénk, ha teljes bizonyossággal kizárhatnánk minden ettől eltérő eredményt. Ugyanakkor az időzítésbeli változatosság csökkenése, a rövid és a hosszú magánhangzók időtartamainak közeledése, amelyekkel kapcsolatban Kassai szintén közölt adatokat, hozzájárulhat a beszéd gyorsulásának érzetéhez.

A probléma újragondolását az is indokolja, hogy az utóbbi években született kutatási eredmények sem erősítették meg egyértelműen a beszédtempó gyorsulását. A gyorsulás tényét alátámasztja Menyhárt (2010), aki a Hegedűs-archívumból vett beszédmintákból, illetve mai városi és falusi beszélők spontán beszédéből mért artikulációs és beszédtempót. Az eredményei szerint a leggyorsabban a mai budapesti beszélők, lassabban a mai falusi beszélők, leglassabban pedig az archívumi falusi személyek beszéltek. Ezekből az adatokból következtetett Menyhárt a beszédtempó gyorsulására. Ugyanakkor Vallent (2005) esettanulmányában azt találta, hogy az 1950-ben készült hangfelvételen a vizsgált személyek valamivel gyorsabban beszéltek, mint a 2005-ös kísérletben. Mivel a kísérleti személyek középkorúak és idősebbek voltak, Vallent arra következtetett, hogy a beszédtempó gyorsulása vélhetően inkább a fiatalabb korosztályra jellemző. De nem mutatkozott eltérés a 60 évvel ezelőtti és a mai gyermekek beszédtempóinak összehasonlítása során sem (Menyhárt 2012). Ha pedig Hegedűs Lajos és Fónagy Iván fent idézett adatait összevetjük Imre Angéla 2005-ben kapott eredményeivel, akkor azt kapjuk, hogy a korábban leggyorsabb műfaj, sportközvetítés tempója a korábbi 14 (Hegedűs), illetve 13,8 (Fónagy–Magdics) hang/s értékről lecsökkent 12,5 hang/s-ra, a hírolvasás tempója az 1957-es 12 hang/s-ról felgyorsult 14,4 hang/s-ra, a versmondás tempója a Fónagy–Magdics által közölt 9,4 hang/s-nál csak valamivel lett magasabb, 9,9 hang/s, a beszélgetés (spontán dialógus) tempója viszont az 1960-ban mért 12,89 hang/s értékhez hasonló, 12,7 hang/s 2005-ben. Ezek az eredmények arra utalnak, hogy a beszédtempó változása műfajonként eltér, azonban ez az összehasonlítás is részben a kis számú kísérleti személy, részben az esetlegesen eltérő módszertan miatt óvatosan értelmezendő. További mai adatokat közöl különféle műfajú szövegek tempóértékeire Markó (2005). Mérései szerint négy kísérleti személynél a szabad narratívák artikulációs tempója 12,45–15,61 hang/s, a „kontrollált” narratíváké 12,38–15,90 hang/s, a képleírásoké 11,88–14,95 hang/s, a társalgásé pedig 14,43–16,44 hang/s volt.

A problémakör újbóli vizsgálatának harmadik oka az, hogy a beszédelemző szoftverek segítségével ma már hatékonyabban tudunk méréseket végezni, mint 50-60 évvel ezelőtt, továbbá lényegesen könnyebben tudunk az eddigi-eknél jóval megbízhatóbb eredményeket szolgáltató matematikai statisztikai számításokat végezni a kapott mérési adatokon. Ez pedig felveti annak szükségességét, hogy sok évtizeddel ezelőtti hangfelvételeken ismét vagy akár első alkalommal méréseket végezzünk, és a beszélő neve, életkora, illetve a beszéd műfaja szerint differenciálva még árnyaltabb képet kapjunk a beszéd akusztikai szerkezetének diakrón változásairól.

Jelen kutatásunkban arra a kérdésre keressük a választ, hogy a nyolc évtizeddel ezelőtti filmszínészi beszédben milyen tempóértékeket találunk. A filmszínészi beszéd akusztikai-fonetikai vizsgálata a kevésbé kutatott területek közé tartozik. Nemzetközi vonatkozásban megemlíthető Nitta és munka-

társai (2010) tanulmánya, amelyben a szerzők 14, 1954 és 2003 között készült amerikai film tempóértékeit közlik. A vizsgált filmek tempóértékeinek mediánja összesítve 5,1 szótag/másodperc volt, míg az egyes filmek tempóinak mediánjai 4,3 és 5,8 szótag/másodperc közé estek. A szerzők nem találtak műfajra jellemző különbségeket, egyedül az animációs filmeknél találtak a többinél lassabb tempókat. A régebbi filmekben mért tempók sem voltak lassabbak a maiaknál. Egy másik kutatás szerzői az *Ally McBeal* című amerikai televíziós sorozat férfi és női főszereplőjének eredetiben felvett beszédét, illetve a sorozat német és japán nyelvű, szinkronizált változatát elemezték (Braun–Oka 2007). Az eredeti felvételeken a női főszereplő 5–6, a férfi főszereplő pedig 4,5–6 szótag/másodperces tempókkal beszélt a különböző érzelmi állapotokban. Hasonló értékek adódtak a német szinkron esetében, azaz a kivétellel, hogy a férfi szereplő német szinkronhangja semleges érzelmi állapotban több mint 8 szótag/másodperces tempóval beszélt. A japán szinkronszínészek viszont lényegesen gyorsabban artikuláltak, a női szereplőnél 6–9 szótag/másodperc, a férfinéél pedig kb. 8–9 szótag/másodperc adódott.

Magyar filmekre vonatkozóan egyedül Mártonfi kutatása ismeretes, melyről a szerző a 2007-ben tartott Beszédkutatás konferencián számolt be. Kérdésfeltevése hasonlít jelen tanulmányunkéhoz, azzal a különbséggel, hogy két film eredeti és újraforgatott változatában, a filmek egészére vonatkoztatott artikulációstempó-értékeket vizsgálta. A *Hyppolit, a lakáj* című 1931-ben készült filmben 12,4 hang/s-ot, az újraforgatott, *Hippolyt* című változatában (1999) 12,2 hang/s-ot mért, míg a *Meseautó* eredeti változatában (1934) 13,5 hang/s-ot, a 2000-ben újraforgatott változatában pedig 13,1 hang/s-ot kapott. Ezekből az adatokból kézenfekvő lehet arra következtetni, hogy – legalábbis a filmek tükrében – beszédünk tempója nem gyorsult hat-hét évtized alatt. Ha azonban a filmek egészére vonatkozóan számítjuk ki a tempóértékeket, a főszereplők mondatai szükségképp nagyobb arányban lesznek a vizsgált mintában, így a végeredményben is nagyobb súllyal szerepelnek (vö. Nitta et al. 2010). Ez nem okozna problémát, azonban ha ugyanazt a szerepet két eltérő karakterű, életkorú, művészi felfogású színész alakítja, a művészek egyéni jellemzői is nagyban differenciálhatják színészi beszédük temporális szerkezetét. Így – ha a film egészére vonatkozóan mérünk tempót – ugyanazon film régi és új változatában mért tempókülönbségek nem kizárólag annak lesznek tulajdoníthatók, hogy mikor készültek, hanem annak is, hogy kik és milyen felfogásban játszották az adott szerepeket, illetve milyen arányban szerepelt beszédük a filmben. Célszerűbbnek és megbízhatóbbnak tűnik tehát külön-külön, személyenként vizsgálni a filmekben szereplő színészek beszédét. Jelen munkánk ezt a feladatot tűzi ki célul. Hipotézisünk szerint az 1930-as évek filmjeiben a színészek a mainál alacsonyabb tempókkal beszéltek.

Anyag és módszer

A kutatáshoz a *Pesti mese* és a *Fizessen, nagysád!* című, 1937-ben készült magyar játékfilmeket használtuk fel, amelyek műsoros DVD-n álltak rendelkezésre. A filmekben két-két főszereplő, Turay Ida (30 éves), Páger Antal (38 éves), illetve Muráti Lili (25 éves) és Jávor Pál (35 éves) beszédét vizsgáltuk. Az artikulációs tempó méréséhez a filmek hangsávját a Praat 5.3.55 programba importáltuk (Boersma–Weenink 2013). A mérés alapegysége a társalgási helyzetekben elhangzott mondat volt. Ezeket könnyen lehetett azonosítani, mivel a színészek a szövegüket előre megírt forgatókönyv alapján mondták. Célként tűztük ki, hogy a méréshez a kiválasztott személyek által a két filmben kimondott valamennyi mondatot felhasználjuk, azonban erre – elsősorban technikai okok miatt – nem volt lehetőség. A mérésekhez olyan mondatokat használtunk fel, amelyek:

- nem tartalmaztak (a vélhetően filmszakadás miatti) ugrásokat, kihagyásokat
- nem tartalmaztak olyan részeket, amelyekben két szereplő egyszerre beszélt, és nem lehetett elkülöníteni a hangjukat
- nem tartalmaztak a mérés szempontjából zavaró háttérzenét, háttérzajt
- elhangzásukkor a társalgásban részt vevő személyek egyszerre jelen voltak, azaz telefonbeszélgetést nem vizsgáltunk
- tiszta artikulációval jelentek meg, azaz a szereplő nem evett vagy nem torzította el az arcát, nem mutatkozott betegnek, ittasnak stb. beszéd közben.

Emellett az általában izoláltan ejtett, egyszavas kifejezéseket sem vizsgáltuk, ezek a legtöbbször erős érzelmi töltettel, elnyújtottan ejtett indulatszavak voltak, számuk azonban nem volt jelentős. A korpuszba felvett mondatokat fajtájuk, érzelmi töltetük stb. szerint nem differenciáltuk, ez további vizsgálatok tárgya lehet.

A fenti feltételeknek megfelelően a kutatás során Turay Ida 120, Páger Antal 110, Muráti Lili 115, Jávor Pál 94 mondatát használtuk fel. Az artikulációs tempó kiszámításához a Praat 5.3.55 programban megmértük az egyes mondatok hosszát, majd megszámláltuk a kiejtett beszédhangokat. A kapott értékeket az SPSS 13.0 programban rögzítettük, majd kiszámoltuk az artikulációs tempókat. Az időtartamok méréséhez az egyes mondatok oszcillogramján kívül megjelenítettük a hangszínképüket is, ami megkönnyítette a mondathatárok megállapítását. A hipotézis igazolásához mindegyik beszélőre vonatkozóan gyakorisági eloszlási grafikonokat készítettünk, majd kiszámítottuk az artikulációs tempók átlagát. Az átlagtól való eltérések jellemzésére többféle mennyiség is alkalmazható, például az átlagtól való átlagos eltérés vagy a szóródás, azaz a standard deviáció. Tanulmányunkban ez utóbbit használtuk és SD-vel jelöljük. Az eloszlás vizsgálatához a Kolgomorov–Smirnov-próbát, illetve a csúcsosság (kurtosis) és a ferdeség (skewness) értékeit vettük alapul.

A filmekben nem voltak olyan hosszabb szövegrészek, amelyekből megbízhatóan lehetett volna átlagos beszédtempót mérni, így vizsgálatunk csak

az artikulációs tempó számítására korlátozódik. A mondatok túlnyomó többségét a beszélők egy levegővétellel mondták el. Azokban az esetekben, amikor levegővétel vagy szünet szerepelt a mondatban, ezek időtartamát levontuk a mondat teljes hosszából az artikulációs tempó számításához. Néhány esetben egy-egy beszédlepcs (vö. Dér 2012) két-három mondatot is tartalmazott. Ezekben az esetekben külön-külön kezeltük a mondatokat, így az artikulációstempó-értékekről még árnyaltabb képet kaphattunk.

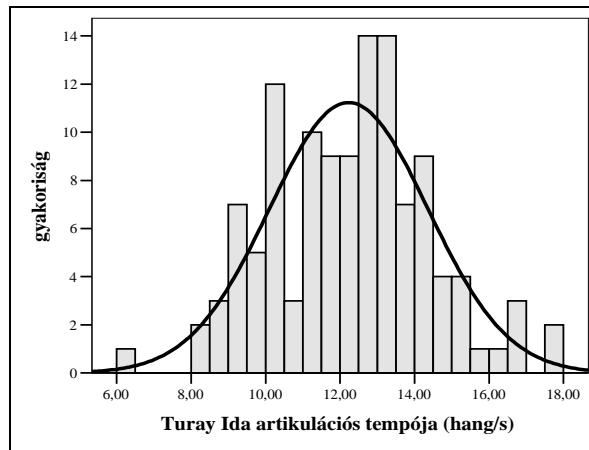
A mérések elvégzésekor egy technikai kérdés merült fel, amely nem hagyható figyelmen kívül az eredmények értékelése során. A filmkamera szokásos sebessége 24 kép/s, videóra történő átíráskor azonban 25 kép/s lesz a lejátszási sebesség (Vagyóczky 1999). Feltételezhető, hogy a másodpercenkénti egy képkockás sebességnövekedést a film kismértékű gyorsításával érték el a digitalizálás során. Eredetileg tehát egy képkocka 0,04167 másodpercig volt látható, a digitális átírást követően azonban, ha feltételezzük a gyorsítást, akkor csak 0,04 másodpercig jelenik meg. Ha a felgyorsított filmből az eredeti időviszonyokat szeretnénk visszakapni, akkor a mért időtartamokat 4,167%-kal meg kell növelni, azaz 1,04167-tel meg kell szorozni. Így a tempóértékek valamivel alacsonyabbak lesznek. Mivel nem áll rendelkezésünkre információ az esetleges gyorsítással kapcsolatban, az eredmények számszerű ismertetésénél mindkét értéket megadjuk, először a ténylegesen mért adatokat, zárójelben pedig a korrigált értékeket.

Eredmények

Az 1. ábra Turay Ida 120 mondatán mért artikulációs tempók gyakorisági eloszlását mutatja be. Az artikulációs tempó átlagos értéke 12,23 (11,74) hang/s, SD = 2,13 (2,044) hang/s. A Kolgomorov–Smirnov-próbat eredménye: $Z = 0,658$, $p = 0,779$, ami arra utal, hogy az eloszlás nem különbözik szignifikánsan a normál eloszlástól. Az eloszlás ferdesége $-0,235$, csúcsossága pedig $-0,036$. Ezek az adatok megerősítik az eloszlás normál jellegét.

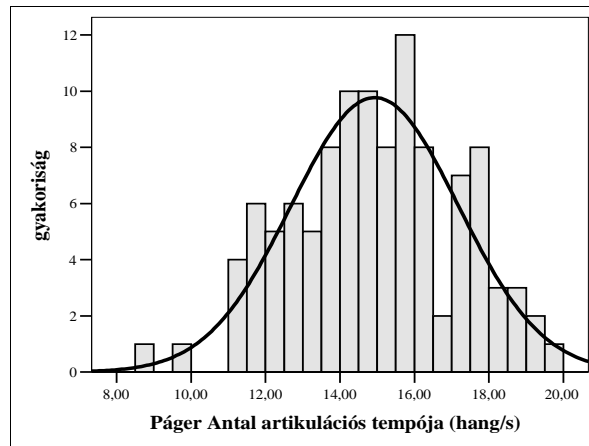
A 2. ábrán látható a Páger Antal 110 mondatán mért artikulációs tempók gyakorisági eloszlása. Az alábbi átlagos artikulációs tempót kaptuk: 14,93 (14,33) hang/s, SD = 2,24 (2,15) hang/s. A normalitásteszt eredménye: $Z = 0,496$, $p = 0,996$, azaz nem szignifikáns az eltérés a normál eloszlástól. A csúcsosság értéke $-0,381$, a ferdeségé $-0,268$. Ezek a 0-hoz közeli értékek megerősítik, hogy normál eloszlást kaptunk.

A 3. ábra Muráti Lili 115 mondatán mért artikulációs tempók eloszlását mutatja be. Átlagos artikulációs tempóra 14,38 (13,81) hang/s, SD = 2,67 (2,56) hang/s adódott. A normalitásteszt eredménye: $Z = 0,965$, $p = 0,309$, ami arra utal, hogy nem szignifikáns az eltérés a normál eloszlástól. Csúcsosságra $-0,587$, ferdeségre $-0,185$ adódik, ezek szintén 0-hoz közeli, a normál eloszlást megerősítő értékek.



1. ábra

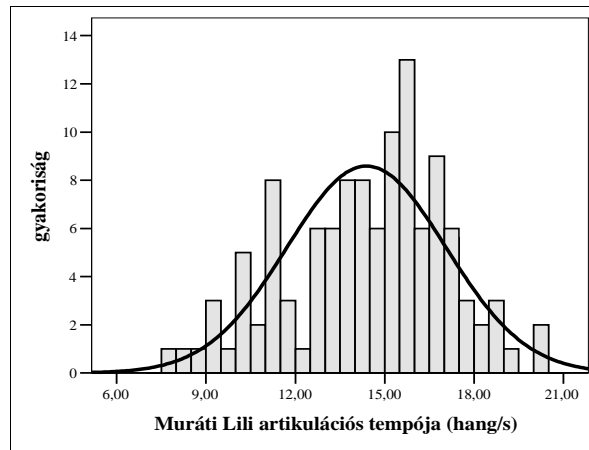
A Turay Ida mondatain mért artikulációs tempók gyakorisági eloszlása



2. ábra

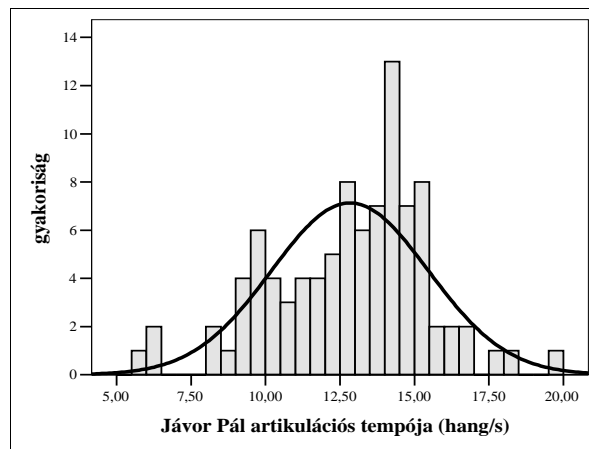
A Páger Antal mondatain mért artikulációs tempók gyakorisági eloszlása

A 4. ábra Jávör Pál 94 mondatán mért artikulációs tempókat összesíti. Az átlagos artikulációs tempó: 12,83 (12,32) hang/s, $SD = 2,63$ (2,52) hang/s. A normalitásteszt eredménye: $Z = 0,853$, $p = 0,460$, azaz nem szignifikáns az eltérés a normál eloszlástól. A csúcsosság értéke 0,297, a ferdeségé $-0,379$, ezek a 0-hoz közeli értékek szintén megerősítik az eloszlás normál jellegét.



3. ábra

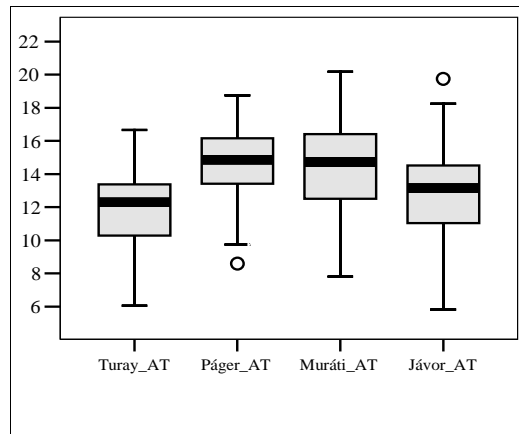
A Muráti Lili mondatain mért artikulációs tempók gyakorisági eloszlása



4. ábra

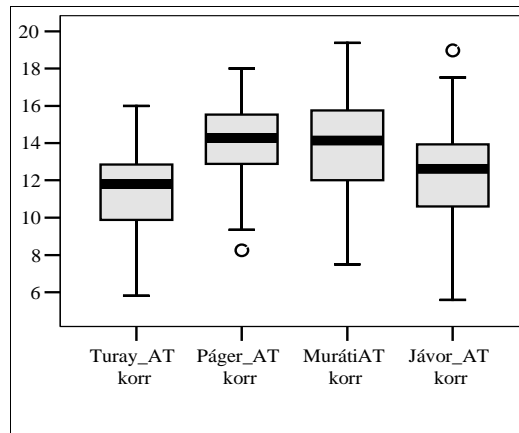
A Jávor Pál mondatain mért artikulációs tempók gyakorisági eloszlása

Az 5. és a 6. ábrán látható a mért és a korrigált artikulációs tempók szóródása dobozdiagramokon. A minimum- és maximumértékek az alábbiak. Turay Ida: 6,06–17,87 (5,82–17,15) hang/s, Páger Antal: 8,59–20 (8,25–19,20) hang/s, Muráti Lili: 7,81–20,18 (7,5–19,39) hang/s, Jávor Pál: 5,83–19,75 (5,59–18,96) hang/s.



5. ábra

Az artikulációs tempók szóródása



6. ábra

A korrigált artikulációs tempók szóródása

Következtetések

Hipotézisünkben feltételeztük, hogy a 76 évvel ezelőtti filmek szereplői lassabban beszéltek, mint amit ma általában mérhetünk. Ha a Bóna (2008) által a közepes beszédtempójú felnőtt beszélők átlagos artikulációs tempójára megadott 12,5–14 hang/s intervallumot vesszük alapul, akkor kijelenthetjük, hogy kutatásunk kísérleti személyei közül Páger Antal és Muráti Lili eseté-

ben ennél gyorsabb, Turay Ida esetében pedig ennél valamivel lassabb átlagos artikulációs tempót kaptunk. A kísérletünk során mért tempóértékek tehát nem alacsonyabbak, mint amelyet mai beszélőknél várnánk. Ha a korrigált értékeket vizsgáljuk, akkor Turay Idánál a jelzett intervallum alatti, 11,74 hang/s-os értéket kapjuk, de még ez sem elképzelhetetlen artikulációs tempó napjainkban. Korábbi kutatásunkban is találtunk olyan mai egyetemista női beszélőt, akinek hasonló volt az artikulációs tempója (Gocsál 2001). A korrigált átlagos tempóértékek Jávor Pál és Muráti Lili esetében a 12,5–14 hang/s közötti intervallumba estek, míg ez az érték Páger Antal esetében valamivel magasabb, 14,33 hang/s volt. Hipotézisünk tehát nem igazolódott, a vizsgált négy beszélő mindegyike olyan artikulációs tempót használt, amely ma is elfogadható, sőt több esetben 18 hang/s-nál gyorsabb tempót mértünk, amely ma is kifejezetten gyorsnak minősül.

Mi lehet ennek az oka? A kapott adatok alapján kézenfekvőnek tűnhet annak megállapítása, hogy beszédünk korábbi eredményeknek ellentmondva nem gyorsult fel a 76 évvel ezelőtti tempóhoz képest. Ám ha a bevezetőben említetteknek megfelelő, beszélők és műfajok szerinti differenciáló szemléletmódot alkalmazzuk, akkor csak azt mondhatjuk, hogy a kutatásban szereplő (i) 25–38 éves személyek, akik (ii) színészek, azaz professzionális beszélők (iii) egy előadóművészi műfajban, azaz reprodukív beszédükben (vö. Imre 2005) produkáltak maihoz hasonló tempóértékeket. Így ebből még nem következtethetünk arra, hogy az előadóművészi, beszédtechnikai szempontból nem képzett köznyelvi beszélők a hétköznapi spontán beszédükben is ilyen tempókat alkalmaztak volna. Csak azt jelenthetjük ki, hogy találtunk egy olyan beszédműfajt, amelyben 76 évvel ezelőtt előforduhattak a maihoz hasonló tempóértékek. Lehetséges, hogy ez a tempó műfaji sajátosság volt, és a színészek az intenzívebb hatáskeltés miatt használták. De az is elképzelhető, hogy életkoruk, illetve városi életmódjuk tükröződött artikulációs tempójukban. Ugyanakkor, mivel ezek a mondatok mozifilmekben hangoztak el, a korabeli köznyelvi beszélők tömegesen tapasztalhatták meg az itt mért, a maihoz hasonló tempóértékeket.

Kutatásunk további eredménye, hogy ugyanazon személynél különböző társalgási helyzetekben mért artikulációs tempók – legalábbis a vizsgált előadóművészi produkciókban – normál eloszlást mutatnak. Ez azt jelenti, hogy van egy kényelmes átlagos tempó, amely körül a normál eloszlásnak megfelelően szóródnak az egyes mondatoknál mért tempóértékek. Megfigyelhető ugyanakkor, hogy mindegyik esetben negatív előjelű lett a ferdeség értéke. Ez, bár 0-hoz volt közel, arra utal, hogy az eloszlási grafikonok maximumértéke alatt valamivel nagyobb tartományban szóródtak a tempóértékek, mint fölötté. Bár a normál eloszlástól nem volt szignifikáns az eltérés egyik esetben sem, a maximumérték enyhén jobbra való eltolódása tendenciaszerűen mindegyik grafikonon megfigyelhető.

További kutatási feladat annak vizsgálata, hogy melyek azok a tényezők, amelyek a tempóértékek szóródását okozzák. Feltételezhető, hogy a beszédhelyzet, a beszélő érzelmi állapota, a beszélgetőpartner személye, a beszélgetés tárgya egyaránt szerepet játszhat ebben. Szintén további kutatást igényel az is, hogy ugyanezeknek a színészeknek a spontán beszéde milyen tempót mutatott, illetve az is, hogy vajon egy mai professzionális vagy hétköznapi beszélő esetében is hasonlóan szóródnak-e a tempóértékek, ha például egy személy teljes napi beszédprodukciónak vizsgálunk hasonló módszerrel.

Eredményeinkből – mivel csak négy beszélőt vizsgáltunk – általános következtetést nem vonhatunk le. Mai filmekkel ugyan nem végeztünk összehasonlítást, így csak feltételezésként fogalmazzuk meg, hogy a kapott adatok tükrében a fiatal színészek filmszínészi beszéde nem gyorsult nyolc évtized alatt.

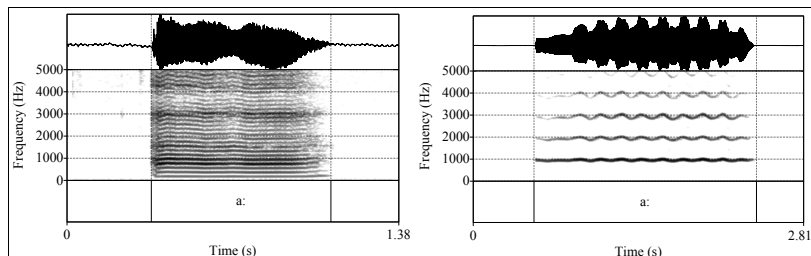
Irodalom

- Boersma, Paul – Weenink, David 2013. *Praat: doing phonetics by computer [Computer program]. Version 5.3.* <http://www.praat.org>. (Letöltés ideje: 2013. júl. 9.)
- Bóna Judit 2008. A beszédtempó pedagógiai vonatkozásai. *Anyanyelv-pedagógia* 2008/1. <http://www.anyanyelv-pedagogia.hu/cikkek.php?id=16>.
- Bóna Judit 2010. Bizonytalansági megakadások idősek és fiatalok beszédében. *Beszédkutatás* 2010. 125–138
- Braun, Angelika – Oba, Reiko 2007. Speaking tempo in emotional speech – a cross-cultural study using dubbed speech. In: *Proceedings of the International Workshop on Paralinguistic Speech - between models and data, ParaLing '07*, 3 August 2007, Saarbrücken, Germany. <http://www2.dfki.de/paraling07/papers/16.pdf>. (Letöltés ideje: 2014. február 8.)
- Dér Csilla Ilona 2012. Beszélőváltások során használt diskurzusjelölők a magyar spontán beszédben. *Beszédkutatás* 2012. 132–143.
- Gocsál Ákos 2001. Gyorsabban beszélnek-e a nők, mint a férfiak? *Beszédkutatás* 2001. 61–72.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- Gósy Mária – Bóna Judit – Beke András – Horváth Viktória 2013. A kitöltött szünetek fonetikai sajátosságai az életkor függvényében. *Beszédkutatás* 2013. 121–143.
- Horváth Viktória 2007. Vannak-e „női” és „férfi” megakadásjelenségek a spontán beszédben? *Magyar Nyelvőr* 131. 315–323.
- Imre Angéla 2005. Különböző műfajú szövegek szupraszegmentális jellemzői. *Magyar Nyelvőr* 129. 510–520.
- Kassai Ilona 1993. Gyorsult-e a magyar beszéd tempója az elmúlt 100-120 évben? *Beszédkutatás* 1993. 62–69.
- Markó Alexandra 2005. *A spontán beszéd néhány szupraszegmentális jellegzetessége, monologikus és dialogikus szövegek összevetése, valamint a hűmmögés vizsgálata*. Doktori értekezés. ELTE, Budapest. <http://spontanbeszed.hu/letoltes/asponanbeszed-szuprasz.pdf>. (Letöltés ideje: 2014. február 10.)
- Mártonfi Attila 2007. Beszédtempó-változás a filmremake-ek tükrében. Elhangzott: Beszédkutatás 2007 konferencia.

- Menyhárt Krisztina 2010. A beszédsebesség objektív és szubjektív észlelései eredményeinek összefüggései mai és 60 évvel ezelőtti beszélőknél. *Beszéd kutatás 2010*. 110–124.
- Menyhárt Krisztina 2012. A beszéd temporális jellemzői 60 évvel ezelőtti gyermek beszélőknél. *Beszéd kutatás 2012*. 246–259.
- Nitta, Haruhiko – Okazaki, Hironobu – Klinger, Walter 2010. An analysis of articulation rates in movies. *ATEM Journal, The Association for Teaching English Through Movies* 15. 41–56. <http://www.office.usp.ac.jp/~klinger.w/2010-An-Analysis-of-Articulation-Rates-in-Movies.pdf>. (Letöltés ideje: 2014. február 8.)
- Vagyóczky Tibor 1999. Alapismeretek az operatőri munkához. In Vagyóczky Tibor (szerk.): *Kézikönyv film- és tévéalkotóknak*. Magyar Operatőrök Társasága, Budapest. 11–31.
- Vallent Brigitta 2005. A spontán beszéd ötven éve és ma. (Esettanulmány). *Beszéd kutatás 2005*. 99–111.

Az operatőreklésben elvárás, hogy az énekhang szinte a teljes alaphangtartományon homogén hangszínezetű és nagy intenzitású legyen. Ezt az énekesek – a magasabb alaphangokon – olyan artikulációs gesztusok segítségével érik el, melyek sok esetben nem kedveznek a magánhangzókra a mindennapi beszédben jellemző artikulációs konfigurációnak. A magánhangzók percepciója szempontjából is talán legfontosabb ilyen gesztus az állkapocs nyitáshoz vezető alaphang emelésével párhuzamos növelése. Ez a technika főként azon alaphangmagasságok esetében használatos, melyek esetében igaz az, hogy az alaphang (f_0) frekvenciaértéke nagyobb, mint az ejteni kívánt magánhangzó a beszédben megjelenő első formánsának (F_1) értéke. Ezek a hangmagasságokon ugyanis az artikuláció megváltoztatása nélküli ejtés során az F_1 nem jelenhetne meg a magánhangzó spektrumán, és az alaphang intenzitása is alacsonyabb, következésképpen az egész spektrum energiája is kisebb lenne. Az állkapocs nyitáshoz vezető növelése azonban megemeli az F_1 értékét, így a fent említett helyzetben a magas frekvenciaértékű f_0 -nál alacsonyabb frekvenciaértékű F_1 feljebb (az f_0 értékére vagy akár afölé is) hangolható (Sundberg 1987; Garnier et al. 2010). Artikulációs tekintetben tehát az állkapocs nyitáshoz vezető, míg akusztikai tekintetben az F_1 az alaphang függvényében változik – legalábbis azokban az esetekben, amikor az alaphang eléri a beszédben megjelenő F_1 értékét.

Ám az énekelt magánhangzók akusztikai szerkezetét nemcsak az artikulációs változások, hanem a magas f_0 artikulációs mozgásoktól független akusztikai sajátosságai is befolyásolják. A mélyebb és magasabb alaphangon ejtett magánhangzók a magas f_0 -ra visszavezethető akusztikai különbségét szemlélteti az 1. ábra, melyen az [a:] hang keskeny sávú spektrogramja látható kétféle ejtésben: a beszédben ($f_0 \sim 190$ Hz) (bal oldal) és a B5 zenei hangon ($f_0 = 988$ Hz) énekelve (jobb oldal), egy szoprán énekesnő ejtésében.



1. ábra

Balra: az [a:] hang keskeny sávú spektrogramja a beszédben ($f_0 \sim 190$ Hz);
 jobbra: az [a:] hang keskenysávú spektrogramja a B5 énekelt
 alaphangmagasságon ($f_0 = 988$ Hz)

A keskeny sávú spektrogramon látható vízszintes csíkok a felharmonikusok, melyek frekvenciaértékei az alaphang egész számú többszörösei. A spektrogramok szemléltetik, hogy a magasabb alaphangú hang esetében a felharmonikusok (az alaphanggal való összefüggésük következtében) jóval távolabb esnek egymástól, mint a mélyebb alaphangú vokálisoknál. A spektrografikus megjelenítésben a frekvenciakomponensek intenzitása is megjelenik, melyet a sötétítés mértéke mutat. Ez jól láthatóan eltér az ábrázolt beszédhangok különböző felharmonikusai esetében – annak megfelelően, hogy a felharmonikusok beleesnek-e valamely formáns (azaz a toldalékcső mint rezonátorüreg valamely sajátfrekvenciájának) sáv szélességébe. A spektrogram kellően mély alaphang és sűrű felhangszerkezet esetén (1. ábra, első kép) tehát a felhangok intenzitásának különbségein keresztül (sötétebb sávokként) „kirajzolja” a formánsokat (itt a beszéd esetében ezek 1 kHz, 1,5 kHz és 3 kHz környékén láthatók), míg erre a magas alaphangú hang felhangszerkezete (1. ábra, második kép) csak jóval kisebb mértékben képes. A magas f_0 következtében egymástól távolabb eső felharmonikusok ugyanis spektrálisan alul-mintavételezik (*undersampling*, l. Goldstein 1980: 235) a toldalékcső átviteli karakterisztikáját, azaz a toldalékcső artikulációs konfigurációját jellemző függvényt, illetve a magánhangzóspektrumot (l. pl. Goldstein 1980; Ryalls–Lieberman 1982). Ennek eredményeként a magas alaphangon produkált énekhang akusztikai kimenete sokkal kevesebb információt

közvetít az artikulációs konfigurációról, azaz az ejteni kívánt beszédhang képzésére berendezkedő beszédképző szervek aktuális állásáról, mint a mélyebb alaphangokon énekelt vagy beszélt vokálisoké.

Az énekhang fent említett tulajdonságaiból látszik, hogy az énekelt magánhangzók ejtése és akusztikai szerkezete szoros kapcsolatban áll az f_0 értékével. Ezzel összefüggésben pedig az a feltételezés is megalapozottnak tűnik, hogy az f_0 értéke az énekelt magánhangzók észlelésében is fontos szerepet játszik. Ennek a feltételezésnek megfelelően az énekelt magánhangzók észlelését vizsgáló szakirodalomban számos kutatás fókuszál az f_0 hatására. Ezek a kutatások kimutatták, hogy az azonosítás sikeressége az f_0 emelésével csökken (Scotto di Carlo–Germain 1985; Gottfried–Chew 1986; Hollien et al. 2000; Evgrafova–Evdokimova 2012; Deme 2012). Ugyanakkor az adatok arra is fényt derítenek, hogy a jól látható csökkenés ellenére bizonyos hangzók bizonyos körülmények között még magasabb alaphangon is felismerhetők maradnak, továbbá arra is, hogy az azonosítás sikerességének csökkenése sem minden esetben mutat szigorúan csökkenő tendenciát (Scotto di Carlo–Germain 1985; Hollien et al 2000; Deme 2012). Ennek ellenére kevés az olyan jellegű kutatás, amely a vokálisok percepciók stabilitását támogató paramétereket, például a mássalhangzó-környezetben rejlő (dinamikus) akusztikai kulcsokat vizsgálná – a legtöbbször csak a magánhangzók ejtési sajátosságainak és az éneklés artikulációs kívánalmainak egybevetésével vagy eltéréseivel magyarázzák a felismerési arányok alakulását.

A beszédészlelés szakirodalmában széles körben elfogadott az a feltételezés, hogy könnyebb a magánhangzók azonosítása akkor, ha mássalhangzó-környezetben halljuk őket. Strange és Verbrugge (1976) híressé vált kísérletében /pVp/ szekvenciákban tesztelte a mássalhangzó-környezetben ejtett magánhangzók azonosítását, és azt a következtetést vonta le, hogy a hang-környezet pozitív hatása a mássalhangzó és magánhangzó határán álló koartikulációs hangátmenetekben kódolt információban rejlik. Bár a vizsgálat eredményei egyértelműen csak a beszélve ejtett magánhangzókra vonatkoznak, mégis előfordul, hogy a magas alaphangon énekelt beszédhangok észlelhetőségével kapcsolatban is említik a szerzők az észlelést valószínűsíthetően segítő mankók számbavételekor a mássalhangzó-környezetet (l. pl. Scotto di Carlo–Germain 1985; Hollien et al 2000). Teszik ezt annak ellenére, hogy a fentebb említett, az énekelt és beszélt ejtés közti jelentős artikulációs és akusztikai eltérések az éneklés kutatóinak körében köztudottak. Azon feltételezés tudományos igazolására, hogy a mássalhangzó-környezet megléte vagy hiánya valóban hatással van-e a magas alaphangon (szoprán éneklésben) ejtett vokálisok azonosítására, eddig összesen egyetlen kutatás született, egy további pedig mélyebb alaphangokon, egy férfi hangkategóriában (egy tenor énekes ejtésében) vizsgálta a kérdést.

Smith és Scott (1980) a magas alaphangon énekelt magánhangzók azonosítását /bVd/ hordozóhangsorban vizsgálta egy szoprán énekes ejtésében. Per-

cepciók adataik a szerzők értelmezése szerint az operaéneklésre nézve is megerősítik azt a beszédben tapasztalható tendenciát, miszerint a szomszédos mássalhangzók segítik a magánhangzók felismerését. Adataikat azonban nem szigorúan kontrollált körülmények között nyerték, hiszen a magánhangzót közrefogó mássalhangzók képzéshelye eltérő volt, így sem azt nem tudni, valóban mely mássalhangzó befolyását tesztelték, sem azt, hogy van-e különbség a magánhangzót megelőző és követő hangátmenetek hatása között. Az sem egyértelmű, hogy milyen módon kontrollálták a szerzők az énekes gégejének függőleges helyzetét, mely állításuk szerint szintén fontos befolyásoló tényezőnek bizonyult. Végül pedig a kapott eredmények interpretációjával kapcsolatban is kételyeink ébredhetnek, hiszen valóban egyértelmű tendenciákat csak az F5 zenei alaphangmagasság (azaz 698 Hz) fölött látunk.

Mélyebb alaphangokon (egy tenor énekprodukcijában) ejtett vokálisok azonosításával kapcsolatban Gottfried és Chew (1986) végzett (szintén /bVd/ hangsorokat tesztelő) vizsgálatot. Eredményeik (Smith és Scott 1980 eredményeihez hasonlóan) a mássalhangzó-környezet pozitív hatását mutatták. Am ezzel a kutatással kapcsolatban is felmerülnek a már korábban is említett módszertani (ti. a hangkörnyezet kiválasztásával kapcsolatos) problémák, illetve további kérdések is, melyek az eredmények értelmezését nehezítik. Gottfried és Chew ugyanis az izolált magánhangzók felismerését a hordozóhangsorban ejtett magánhangzók tiszta fázisából kivágott 200 ms hosszúságú hangszeleten tesztelték, nem valóban izoláltan ejtett hangokon, így nem állítható, hogy ennek a kísérletnek az eredményei az izoláltan ejtett hangokra is közvetlenül vonatkoztathatóak volnának. Ezen kívül pedig nem tisztázott például az sem, hogy milyen ablakolással vágták ki a tesztelt hangszeleteket, amely paraméter szintén befolyásolhatja a kapott eredményeket.

Megjegyzendő, hogy az említetteken kívül még egy további olyan vizsgálat született, mely néhány, képzésmódban és képzéshelyben eltérő mássalhangzó alkotta fonetikai környezet hatását vizsgálta a magas alaphangon (női ejtésben) énekelt magánhangzók azonosításakor. Egy korábbi kutatásomban (2012) a nazálisok, illetve a zöngés és zöngétlen frikatívák között ejtett vokálisok azonosítását hasonlítottam össze, és nem találtam különbséget a mássalhangzók mentén (a kutatás CVC szekvenciákat vizsgált, melyekben az első és utolsó hang megegyezett, ebben a kutatásban azonban nem vettem össze az izoláltan és a CVC hangsorban ejtett hangzók felismerését).

Smith és Scott (1980) adataival azt is demonstrálta, hogy az izolált ejtésű magánhangzók sem válnak teljesen azonosíthatatlanná, még magasabb alaphangokon sem – ez pedig ellentmondani látszik az alaphanggal összefüggő artikulációs és akusztikai sajátosságok alapján esetlegesen elvárható tendenciáknak. Eredményeiket összehasonlítva Gottfried és Chew adataival pedig az is látszik, hogy ez az izolált ejtésben tapasztalt stabilitás nagyobb mértékű, mint ami a CVC hangorból kivágott tiszta fázisok azonosításánál tapasztalható. A hangszeletek észlelésének sikeressége (Gottfried és Chew kísérleté-

ben) ugyanis még a mélyebb alaphangtartományban sem haladta meg a 30%-ot, míg az izoláltan ejtett vokálisokat (Smith és Scott kutatásában) mélyebb alaphangokon akár több mint 50%-ban is az ejtési szándéknak megfelelően azonosították a hallgatók (és drasztikus csökkenés csak az F5 zenei alaphangon következett be). A két kutatás összevetése nyomán tehát kézenfekvőnek látszik a feltetelezés, hogy az önmagukban ejtett magánhangzók valamelyest dinamikusabb szakasza, mégpedig a zönge indulása is szolgáltathat olyan dinamikus akusztikai információt, melyre a hangazonosításkor észlelésünk támaszkodni tud. Ennek a vizsgálatá az azonban ismereteink szerint még nem történt meg.

A jelen kutatás célja, hogy kibővítse Smith és Scott (1980) korábbi kísérletét (Smith és Scott 1980, valamint Gottfried és Chew 1986 eredményeinek összehasonlítása nyomán) a magánhangzó-azonosítást befolyásoló változók szigorúbb kontrollálása mellett. Célunk nem csak a mássalhangzó-környezet, de a zöngeindulás hatásának tesztelése is. Hipotézisünk szerint a magánhangzók észlelése mássalhangzó-környezetben a legsikeresebb, melynél az izoláltan ejtett hangzók azonosítása valamivel kevésbé pontos, a legrosszabb azonosítási arányokat pedig a zöngeindulásra vonatkozó információt nem tartalmazó vokálisok esetében tapasztalhatunk. Ugyanakkor azt is feltételezzük, hogy a mássalhangzók és a zöngeindulás pozitív hatása az alaphang emelésével gyengül a magas alaphang okozta alul-mintavételezés, azaz a „rossz felbontás” miatt.

Anyag, módszer, kísérleti személyek

A kutatásban egy szoprán énekes énekprodukciónak rögzítettük. Az énekes a magyar nyelv három akusztikai tekintetben egymástól legtávolabb eső magánhangzóját (/a: i: u:/) ejtette izoláltan („V” kondíció) és /bVb/ hangkörnyezetben („CVC” kondíció), hat énekelt alaphangmagasságon (F3 = 175 Hz, B3 = 247 Hz, F4 = 349 Hz, B4 = 494 Hz, F5 = 698 Hz, B5 = 988 Hz) és beszélve (melynek átlagos alaphangmagassága 191 Hz volt). Ezen kívül disztraktor hangsorokat is rögzítettünk ugyanazon énekessel. A disztraktorok további magyar magánhangzókat (ɒ ɛ e: o: ø: y:) tartalmaztak /bVb/ hangkörnyezetben és izolált ejtésben, beszélve, illetve a B4, F4, F5 zenei alaphangokon énekelve. Ezen stimulusok percepció eredményeit nem elemeztük. A kutatásban használt hordozóhangsor megválasztását a következők motiválták. A mássalhangzó-környezet szigorú kontrollálása érdekében olyan CVC hordozóhangsört használtunk, melyben a két mássalhangzó azonos, így az adataink egyértelműen az adott képzéshelyű mássalhangzóra vonatkoztathatóak. Azonban ahhoz, hogy az eredményeinket a Smith és Scott (1980) kísérletében látottakkal is összehasonlíthatóvá tegyük, az általuk használt /bVd/ hangsor egyik mássalhangzóját kellett megtartanunk. Korábbi kutatási tapasztalataink szerint a CVC hangsor első mássalhangzója gondosabban artikulált az éneklésben, mint az utolsó, ezért a /bVd/ hangsor első mássalhangzójának a megtartása mellett döntöttünk. A magánhangzó zöngeindulásának hatását egy

harmadik kondíció előállításával teszteltük: egy hangeditáló program, a Wavesurfer (Sjölander–Beskow 2009) segítségével („fokozatos beúszás” hatás alkalmazásával) eltávolítottuk az izoláltan ejtett hangok elejét („Vágott” kondíció). A beúsztatást 30 ms-os időtartamon, logaritmikus függvény segítségével valósítottuk meg. A felvételeket csendesített szobában rögzítettük 44,1 kHz-es mintavételezéssel, 16 biten.

A hanganyagokat egy, a Praat programban (Boersma–Weenink 2013) írt és futtatott percepció tesztben vizsgáltuk 22 ép hallású felnőtt részvételével. A teszt ($3 \text{ kondíció} \times 3 \text{ magánhangzó} \times 7 \text{ alaphangmagasság} = 63$ stimulust és 30 disztraktort tartalmazott). Minden tesztelő minden stimulust kétszer hallott, véletlenszerű sorrendben, így egy ülésben összesen 156 hangminta hangzott el. A teszt előtt a kísérleti személyek azt az utasítást kapták, hogy azonosítsák az önmagában vagy hangsorban elhangzó magánhangzót, és döntésüket a képernyőn megjelenő kilenc magánhangzó egyikének megjelölésével rögzítsék. A képernyőn a magyar hosszú (vagy hosszan ejthető) magánhangzókat (/v a: e: i: o: ø: u: y:/) jelenítettük meg ortografikus alakban. A tesztelők a stimulusokat binaurálisan, fejhallgatón keresztül hallották, és a számítógéphez csatlakoztatott egérrel kattinthatnak a kiválasztott magánhangzóra. A teszt átlagosan 10-15 percet vett igénybe, mivel azonban a stimulusok sorrendje minden tesztben eltérő volt, nem volt szükséges figyelembe vennünk a fáradás következtében esetlegesen felléphető tesztzhatást.

A válaszokat tévesztési mátrixokban összesítettük minden stimulus-válasz párt alaphangmagasságokként rendezve. Ezen a ponton le kell szögeznünk, hogy a tanulmányban használt „helyes válasz” és „tévesztés” terminusok a (stimuluslistával irányított) ejtési szándék és a válasz közötti egyezésként vagy eltérésként értelmezendők.

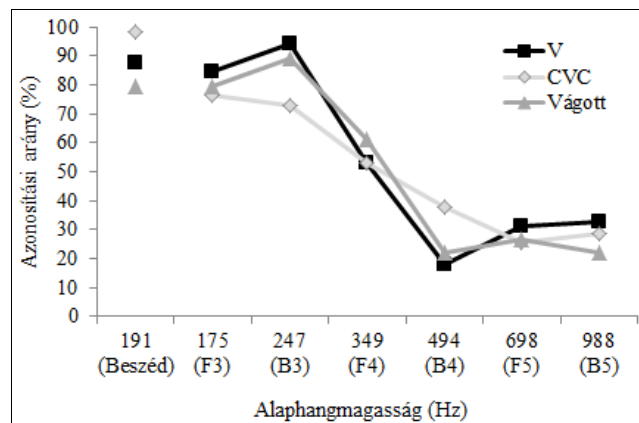
Az adatok statisztikai elemzését az R programmal végeztük (R Core Team 2013).

Eredmények

A szélsőséges teljesítményt nyújtó adatközlők kiszűréséhez minden adatközlőre „konzisztenciaarányt” számoltunk. Ez az érték azt a százalékos arányt mutatja meg, hogy az adott adatközlő az egyes stimulusokat tekintve mennyire volt következetes, azaz mindkét alkalommal ugyanúgy (bár nem feltétlenül helyesen) azonosította-e őket. Mivel minden adatközlő 63 stimulusra adott választ (kétszer), így a következetes válaszok számát (mely tehát maximum 63 lehetett) is erre a számra vetítettük. A konzisztenciaarány arra utal, hogy az adott adatközlőt mennyiben a hangok tényleges azonosítása, és mennyiben csak a találgatás vezette a válaszadáskor. Az első esetben ugyanis a két ismétlés megegyező azonosítása, a második esetben pedig a két ismétlés eltérő felismerése valószínűbb. Ugyanakkor mivel az f_0 emelésével a magánhangzók azonosíthatóságának csökkenését várjuk, ezért az alacsony értékek önmagukban nem az adatközlő megbízhatatlanságát jelzik. Ennek megfelelő-

en pusztán azt vizsgáltuk meg, hogy az értékek egymáshoz viszonyítva egyenletes eloszlásúak-e, azaz homogén csoportot képeznek-e az adatközlők. Mivel a χ^2 -próba tanúsága szerint nincs kilógó érték [$\chi^2(21) = 13,5221$; $p = 0,89$], és az adatok a Shapiro–Wilk-próba szerint normál eloszlásúak, így adatközlőink teljesítményét hasonlóknak ítéltük, és egyetlen adatközlő választait sem kellett kizárunk az elemzésből.

A 2. ábra a három magánhangzó helyes azonosításának százalékos arányát (y tengely) mutatja az alaphangmagasság függvényében (x tengely), összesítve. Az adatok megjelenítése azt sugallja, hogy a vokálisok azonosítása csökken az f_0 emelésével. A kéttényezős ANOVA szerint, melyben az f_0 és a kondíció hatását teszteltük, ez a benyomás helyes, és az f_0 paraméter valóban hatással van az azonosítási arányok alakulására [$F(1) = 33,33$, $p < 0,001$]. A lehetséges összes pár összehasonlítása (Bonferroni-korrekcióval) azt is felfedte, hogy az egymást alaphangmagasság szerint követő énekelt stimulusok között egészen B4-ig (494 Hz) mutatkozik páronként szignifikáns eltérés ($p < 0,02$), mely alaphangtól felfelé azonban a különbségek már nem ilyen jelentősek. A Pearson-féle korrelációs tesztek azt mutatták, hogy az f_0 és az azonosítási arány kapcsolatát viszonylag erős negatív korreláció jellemzi. Amennyiben a három magánhangzóra kapott adatokat egy csoportban vizsgáljuk, erős összefüggést találunk: $r = -0,631$, $p < 0,001$. Ha pedig szétválasztjuk az adatokat a három kondíció mentén, némileg eltérő erősségű, de minden esetben szignifikáns összefüggéseket látunk: V: $r = -0,590$, $p = 0,005$; CVC: $r = -0,609$, $p = 0,003$; vágott: $r = -0,703$, $p < 0,001$.



2. ábra

A három vizsgált magánhangzó azonosításának összesített százalékos aránya az alaphang függvényében

A kéttényezős ANOVA azonban nem igazolta, hogy a kondíció is hatással lenne az észlelési eredményekre (sem önmagában, sem pedig az f_0 változóval kombinálva). A 2. ábrát szemlélve az is kitűnik, hogy a kondíció és az észlelési arányok közötti összefüggés nem is konzisztens, azaz nemcsak az állítható, hogy nem bizonyított a magánhangzók jobb azonosítása mássalhangzó-környezetben, hanem az is, hogy néha (például a B3 zenei alaphangon) kifejezetten rosszabbnak bizonyul a CVC szekvenciában elhangzó vokálisok felismerése. A különbségek statisztikailag ugyan nem igazolhatóak, általánosságban azonban mégis úgy látszik, nagyobb az eltérés a V és CVC, illetve a Vágott és a CVC minták felismerési aránya között, mint a V és Vágott minták esetében.

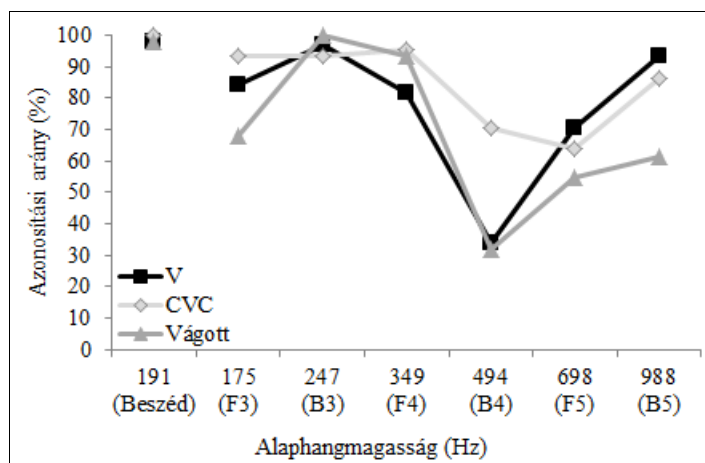
Bár az éneklésben nem igazolódott a fonetikai kontextus (a mássalhangzó-környezet és a zöngéindulás) hatása, a beszédben az elvárt tendenciákat látjuk: a legmagasabb azonosítási arányt a CVC környezet mutatja, melyet az izolált ejtésű, és végül pedig a vágott stimulusok követnek.

Ugyan a fonetikai kontextus hatása a magánhangzó-azonosításra nem igazolódott, az f_0 , a kondíció és a magánhangzó-minőség változókkal lefuttatott varianciaanalízis szerint nemcsak az f_0 , de a magánhangzó-minőség [$F(2) = 22,673$; $p < 0,001$] is, illetve e két faktor kombinációja is [$F(2) = 7,419$; $p = 0,002$] befolyásolja a hangzóazonosítási értékeket. Ezért az azonosítási arányok változását a magánhangzó-minőségek mentén is elemeztük.

A magánhangzó-minőségek közti összes lehetséges páros szembenállás összevetése szerint Bonferroni-korrekciónak után is szignifikáns különbség mutatkozik az /i/-/a:/ és /u/-/a:/ párok, tehát a nyílt és zárt képzésű magánhangzók között ($p < 0,018$), de nincs különbség az állkapocsnyitásszögben egyező ejtésű, zárt vokális, azaz /i:/ és /u:/ között. Ezt az elkülönülést rajzolja ki az f_0 és az azonosítási arány közti korreláció is, ha azt a magánhangzók csoportjain belül vizsgáljuk, ugyanis a két változó csak az /i:/ és /u:/ esetében mutat erős, szignifikáns összefüggést [/i:/: $r = -0,781$ $p < 0,001$; /u:/: $r = -0,900$; $p < 0,001$], míg az /a:/ esetében csak gyenge nem szignifikáns összefüggés látszik ($r = -0,384$; $p = 0,086$). A statisztikai elemzések, továbbá a 3., 4. és 5. ábra tanúsága szerint tehát a zárt képzésű vokálisok azonosításának sikeressége csökken az alaphang emelésével, míg a nyílt képzésű beszédhanghoz tartozó százalékok az f_0 értékétől megközelítőleg függetlenül alakulnak.

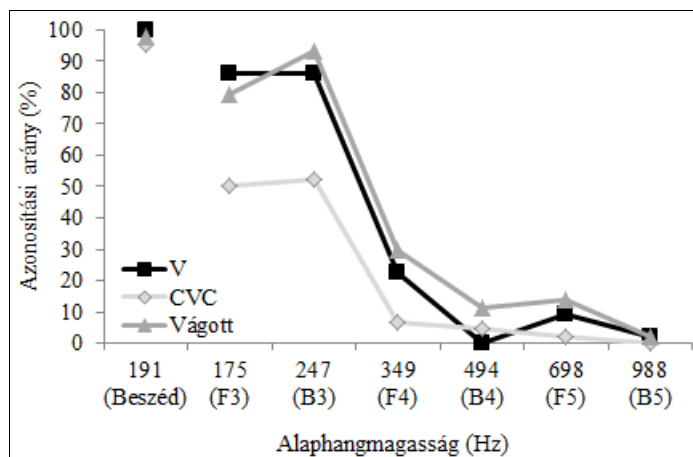
Az /a:/ azonosítási aránya B4 (494 Hz) alatt viszonylag magas minden kondícióban (3. ábra). B4-nél azonban hirtelen esést látunk. Ezen az alaphangon 46%-ban tévesztettek a tesztelők, és minden esetben a magasabb első formánssal ejtett (zártabb képzésű) /v/-t jelölték meg válaszként. A B4 fölött újabb emelkedés következik, majd az azonosítási arányok folyamatosan emelkednek a legmagasabb alaphangig, ahol izolált ejtésben 93%, CVC környezetben 86%, míg eltávolított zöngéindulással 61% a helyes válaszok százalékos aránya. F5-nél ismét az /v/ hang dominálja a tévesztéseket (az összes válasz 27%-a), míg B5-nél az elől képzett /ε/ (az összes válasz 8%-a). Mind-

két hangzó magasabb első formánssal realizálódik beszédben, mint az ejteni kívánt hang.



3. ábra

Az /a:/ azonosításának százalékos aránya az alaphang függvényében

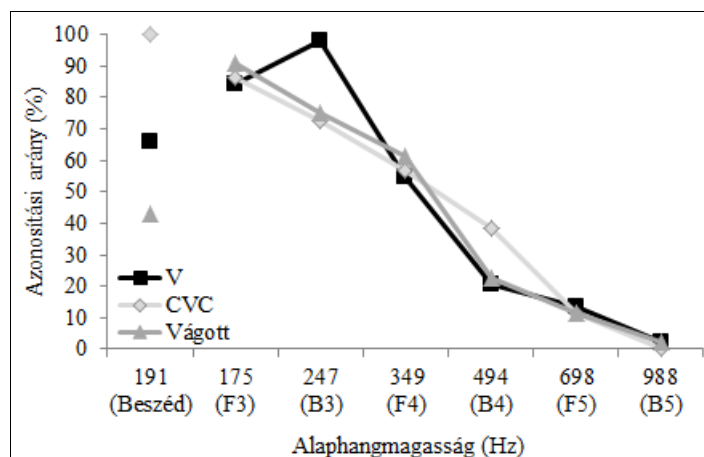


4. ábra

Az /i:/ azonosításának százalékos aránya az alaphang függvényében

Az /i:/ azonosításának százaléka az F4 alaphangnál mutatnak hirtelen csökkenést (4. ábra), itt és a magasabb alaphangok esetében is a hibák legnagyobb része /e/-re történő tévesztés (F4: 80%, B4: 94%, F5: 67% az összes

válaszból). A B5, azaz a legmagasabb alaphang esetében azonban az /i:/-t is leggyakrabban /a:/-ként ismerik fel a hallgatók (az összes válasz 80%-ában), a helyes azonosítás aránya 2% körüli (minden kondícióban).



5. ábra

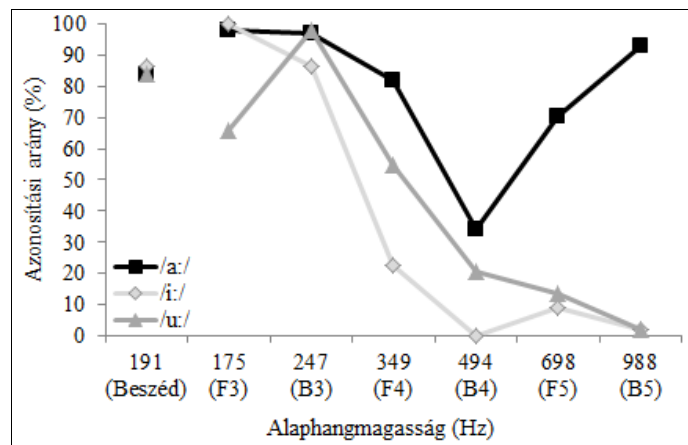
Az /u:/ azonosításának százalékos aránya az alaphang függvényében

Az /u:/ azonosítását leíró görbék (5. ábra) a 2. ábrán bemutatott átlageredményekhez hasonlóan alakulnak: az f_0 emelésével fokozatosan csökkenő százalékarányokat látunk, míg a V és CVC kondíciók között nem következetesen alakuló kapcsolatok mutatkoznak. A hibákat tekintve a legtöbb alaphangmagasságon leggyakrabban előforduló hangzó az /o:/ (F3 és B3: 13%, F4: 42%, B4: 62%, F5: 25% az összes válaszból), ugyanakkor a legmagasabb alaphangon (B5) itt is az /a:/ uralja a tévesztések hierarchiáját (az összes válasz 86%-ában), a vokális azonosítása pedig itt is maximum 2%-ban bizonyult az ejtési szándékkal megegyezőnek.

Az /u:/ beszélt ejtésben bemutatott realizációi (az /i:/-vel és /a:/-val szemben) a hipotézisben megfogalmazottaknak megfelelő tendenciát mutatják: a legjobb arányban a CVC hangsorbéli megvalósulást azonosították a hallgatók, melyet a V és végül a Vágott kondíciók követnek (5. ábra). Azonban ez az /u:/ esetében is csak a beszélt ejtésre korlátozódó mintázat (a többi hang esetében pedig sem a beszédben, sem másutt nem látjuk ezt a tendenciát).

Összehasonlítottuk az összes lehetséges hangzószembenállást a három kondíció alkotta csoporton belül is. Ezek a tesztek megmutatták az egyetlen olyan tendenciát, mely a jelen adatokban a kondícióval, és azon belül is a mássalhangzó-környezet meglétével is összefüggésbe hozhatóan bizonyult. Egyik csoporton belül sem találtuk ugyanis eltérőnek a különböző állkapocs-

nyitásszöggel képzett hangzókat, kivétel a CVC csoportot. Itt ugyanis az eltérő állkapocsnyitással képzett hangok azonosítási arányát eltérő mértékben látszott befolyásolni az f_0 értéke – bár csak a nyelv vízszintes mozgásának tekintetében jobban eltérő pár, az /i:/ és az /a:/ esetében találtunk a Bonferroni-korrekció után is ($p < 0,02$) szignifikáns eltérést (a nyelv vízszintes mozgásának tekintetében egymáshoz hasonlóbb /u:/ és /a:/ nem tért el). Az összehasonlításokat a 6., 7. és 8. ábrán szemléltetjük.



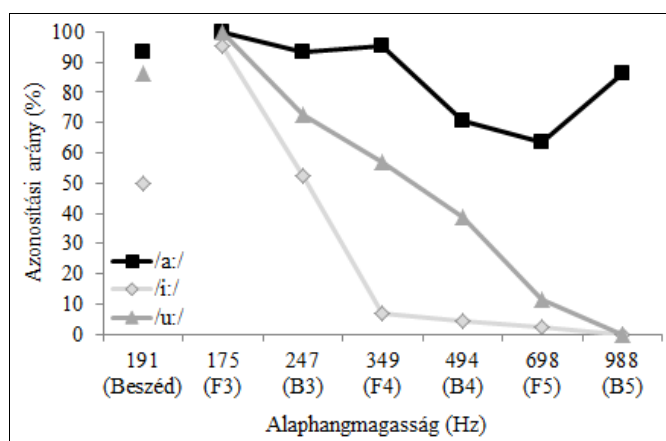
6. ábra

A magánhangzó-azonosítás aránya izolált ejtésben (V kondíció)

Az adatokból jól látható, hogy a különbségek megjelenése (illetve meg nem jelenése a többi kondícióban) annak köszönhető, hogy a CVC kontextusban tapasztalható leginkább az /a:/ percepció stabilitása, ugyanakkor a zárt /i:/ is itt mutatja a legalacsonyabb felismerési arányokat. Míg a V és Vágott kondíciókban csak B4 fölött, addig a CVC kondíciókban már a jóval alacsonyabb F3 fölött is divergáló tendenciákat látunk a három magánhangzóra.

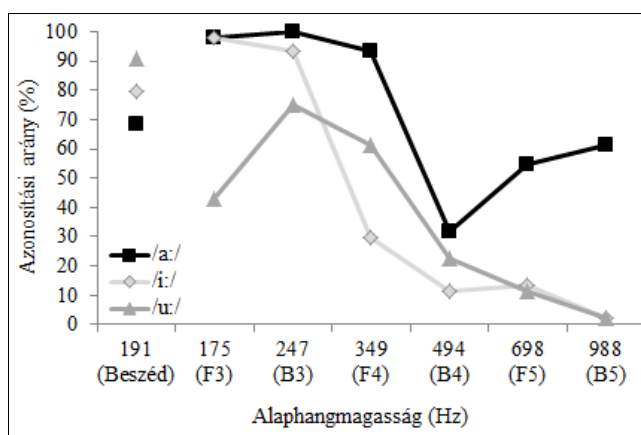
Következtetések

A jelen kutatásban a mássalhangzó-környezetnek és a zöngéindulásnak az énekelt magánhangzók percepciójára kifejtett hatását vizsgáltuk. A két, a témában folytatott korábbi kutatás eredményeire (Smith–Scott 1980 és Gottfried–Chew 1986), valamint a beszédpercepció szakirodalmára (Strange–Verbugge 1976) alapozva azt feltételeztük, hogy a formánsátmenetekben, illetve a zöngéindulásban kódolt dinamikus akusztikai információ segíti a magánhangzók azonosítását – bár ez a hatás az alaphang emelésével (a toldalékcso átviteli karakterisztikájának spektrális alul-mintavételezettségéből fakadóan) csökken.



7. ábra

A magánhangzó-azonosítás aránya mássalhangzó-környezetben (CVC kondíció)



8. ábra

A magánhangzó-azonosítás aránya a zöngéindulás nélkül bemutatott magánhangzók esetében (Vágott kondíció)

A hipotézist egy szoprán énekes ejtésében, három magánhangzó (/a: i: u:/), három kondíció (/bVb/ hangkörnyezet, izolált ejtés, eltávolított zöngéindulás), továbbá 6 énekelt alaphangmagasság (az F3 és B5 zenei hangok között) és a beszélt ejtés segítségével teszteltük egy percepció vizsgálatban, melyben 22 kísérleti személy vett részt.

A kutatás legfontosabb eredménye, hogy sem a mássalhangzó-környezet, sem az izolált ejtésű magánhangzó zöngéindításának pozitív hatását nem sikerült igazolnunk: a kondíciók páronként inkonzisztens eltéréseket mutattak. Az eltérések következtetlensége azt jelenti, hogy nemcsak a CVC kontextus, illetve a zöngéindulás pozitív hatását nem sikerült demonstrálnunk, de éppen az elvárásokkal ellenkező példákat is találtunk (a CVC esetében általánosan a B3 alaphangon, a zöngéindulás esetében pedig a B3 alaphangon az /a:/ hangra). Továbbá mivel a V és Vágott kondíciók gyakorlatilag nem mutattak elkülönülést, ezért azt is feltételezhetjük, hogy a magánhangzó és a mássalhangzók közti formánsátmenetek feltehetőleg nagyobb hatással bírnak az azonosításra, mint a zöngé indulása. Az eredményeket két, egymást kiegészítő tényező kiemelésével magyarázhatjuk: a magas alaphangon ejtett (zöngés) hangok spektrális alul-mintavételezettségével és az operaéneklésnek a mássalhangzók képzését érintő sajátosságaival.

Amint azt a bevezetőben bemutattuk, a magánhangzók spektruma a magasabb énekelt alaphangokon a felhangok nagy távolsága miatt alul-mintavételezetté válik. Azonban ez az alul-mintavételezettség nemcsak a vokálisok tiszta fázisát jellemzi, hanem minden olyan beszédhangot és hangrészletet is, melyek artikulációjának akusztikai lenyomatát a hangszalagműködés következtében létrejövő zöngé hozza létre. Ennek megfelelően a zöngés mássalhangzók és a bármely két szonoráns között realizálódó formánsátmenetek is alul-mintavételezetté válnak. Feltételezhetjük tehát, hogy a kutatás eredményei azt demonstrálták, hogy amiként a magánhangzók tiszta fázisának formánsait, úgy a hangkapcsolatokat átkötő formánsátmeneteket sem kódolja megfelelő pontossággal a magas alaphangú zöngé az énekhang akusztikai lenyomatában, így a formánsmozgásokban rejlő akusztikai kulcsok az éneklés-kor a percepció számára nem elérhetőek.

A másik, az előzőt is kiegészítő magyarázat az énekelt mássalhangzók ejtésén alapszik. Az operaéneklés (vagy másként a *bel canto* technika) egyik alapvetése a magánhangzók időtartamának maximalizálása a mássalhangzók időtartamának rovására. Ennek a mássalhangzó-redukciónak a módját és jellegét azonban, a szerző ismeretei szerint, kísérletes úton eddig még csak egyetlen alkalommal vizsgálták. Deme, Grácsi és Jankovics (2013) az énekelt obstruensek zöngéességének alakulására fókuszáló kutatásában a /b/ zöngés explozíva rövidülését mutatta be szókezdő és szózáró helyzetben, illetve a szókezdő, szó belseji és szózáró explozívak és frikatívák zöngéességének megváltozását is leírják. Ezen túlmenően azonban a szerzők beszámolnak arról is, hogy a mássalhangzók címkézésekor gyakori problémát jelentett az explozívak felpattanásainak elmaradása (vagy legalábbis detektálhatatlansága a spektrogramon és az oscillogramon), illetve a réshangok turbulens zörej-összetevőinek alacsony intenzitása. Az obstruensek pontatlan képzése és egyes akusztikai összetevőinek csekély intenzitása vagy hiánya azonban nemcsak magukra a mássalhangzókra, illetve azok azonosíthatóságára van

hatással, hanem a koartikulációs formánsátmenetekre és azok akusztikai lenyomatára is: ha például a magánhangzót megelőző explozív felpattanása elmarad, a formánsátmenetek sem jelennek meg vagy nem lesznek tisztán ki-vehetőek. Ez pedig nyilvánvaló módon a formánsátmenetek percepció el-érhetőségét is akadályozni fogja.

Mindezek mellett azonban arról sem szabad megfeledkeznünk, hogy ma-gasabb alaphangok esetében a vokálisok artikulációjának jelentős megvá-to-zását, elsősorban nagyobb mértékű állkapocsnyitást kell feltételeznünk. Ha tehát egy megváltozott ejtésű vokális (egy adott alaphangon) percepciósan már nem az ejtési szándéknak megfelelő, hanem egy másik magánhangzó-kategóriába kerül, a koartikulációban rejlő és esetlegesen akár el is érhető in-formáció sem az ejteni kívánt, hanem „csak” erre az utóbbi, az ejteni tudott hangzóra utalhat (azaz valamivel következetesebbé teheti a „tévesztési” ten-denciákat).

Természetesen az adatok kimerítő értelmezéséhez szükséges a különböző képzéshelyű mássalhangzók (tehát nemcsak a bilabiális, de az alveoláris és veláris felpattanók) vizsgálata is, melyek képzéshelyük (illetve eltérő frek-venciaszerkezetük) következtében eltérő formánsátmenetekkel realizálódnak a CVC szekvenciákban. Ezt, tehát a képzéshely okozta különbségeket követ-kező kutatásainkban tervezzük vizsgálni.

A jelen kutatás eredményei ellentmondanak a két korábbi vizsgálatban (Smith–Scott 1980; Gottfried–Chew 1986) olvashatóknak (ahogyan a be-szédpercepció szakirodalma, vö. Strange–Vergbrugge 1976, alapján megfo-galmazható feltételezéseknek is). Kérdés, hogy mivel magyarázható ez az el-lentmondás. Amint arra már a bevezetőben is utaltunk, Smith és Scott ered-ményei az adatokat befolyásoló változók nem szigorú kontrollálása következ-tében nehezen interpretálhatók – ezzel szemben a jelen vizsgálatban mind a felvételek, mind pedig a tesztelés alaposan kontrollált és felügyelt volt. Ezen túlmenően Smith és Scott mindössze 10 kísérleti személy anyagát használta fel – ehhez képest a jelen kutatásban több mint kétszeres adattöménnyiséggel, 22 adatközlő válaszaival dolgoztunk. Harmadrészt feltehető az is, hogy a két kísérletben használt eltérő magánhangzók és azok képzésének (a centralizá-ció mértékének) különbségei is befolyásolták az adatokat. Smith és Scott ugyanis az /i ɪ ɛ æ/ vokálisokat használta, egy feltehetőleg ausztrál akcentusú beszélő ejtésében. Végezetül pedig felmerülhet az a gyanú is, hogy a két ku-tatásban kapott eredmények a kísérletben részt vevő énekesek egyéni ének-technikai sajátosságaiból fakadnak. A szakirodalom alapján azonban jelenleg megválaszolhatatlan az a kérdés, hogy az egyes (profi) operaénekesek ren-delkezhetnek-e olyan mértékben eltérő egyéni artikulációs jellemzőkkel (mind a magán-, mind a mássalhangzók ejtését tekintve), melyek magyaráz-hatnák a korábbi elemzések és a jelen kutatás eredményei közti jelentős elté-réseket, így ennek a feltételezésnek a vizsgálata további kutatásokat igényel. Az mindenestre biztosnak látszik, hogy a kutatások eredményei közötti kü-

lönbség nem a jelen kutatás módszertani felépítésének esetleges problémáiból vagy a jelen adatok kétes hitelességéből fakad.

Bár a mássalhangzó-környezet és a magánhangzók zöngéindulása nem befolyásolta a magánhangzók azonosítását, az f_0 és a magánhangzó-minőség, továbbá a kettő interakciója statisztikailag is igazolható hatást gyakorolt a felismerési arányokra. Ebből következően az elemzés során a magánhangzó-minőségekkel és az alaphangmagassággal összefüggő tendenciákat is részletebben elemeztük. E két változó mentén megközelítőleg a korábban leírtakkal egybevágó tendenciákat találtunk (vö. Scotto di Carlo–Germain 1985; Gottfried–Chew 1986; Sundberg 1994; Deme 2012):

1. Általánosságban elmondható, hogy a magánhangzó-azonosítás feladata egyre nehezebbé vált (azaz a helyes azonosítások aránya csökkent) az alaphang emelésével. A magánhangzó-minőségek mentén azonban eltéréseket tapasztaltunk (lásd 2. és 3. pont).

2. A zárt képzésű /i:/ és /u:/ azonosításának aránya határozottan csökkent az f_0 növelésével párhuzamosan. A beszédben mindkét hangzó első formánása alacsony értéken jelenik meg, melyet tehát már a közepesen magas alaphangok esetében is elér, tehát befolyásol(hat) az f_0 . Mindkét vokálist elsősorban (egy fokkal) nyíltabb, magasabb első formánssal rendelkező hangzókra tévesztették a hallgatók. Az /i:/-re és /u:/-ra kapott percepciós adatok tehát megfelelnek az artikulációs és akusztikai jellemzők alapján feltételezhetőeknek, melyek szerint az f_0 emelése az éneklésben nagyobb állkapocsnyitással, így az első formáns emelésével jár (az eredeti hangzó artikulációs jellemzőitől független módon), a hangzókat pedig ennek megfelelően nyíltabbként is észleljük.

3. A legnyíltabb magyar magánhangzó, az /a:/ azonosíthatósága (a zárt hangzókhoz szemben) még magasabb alaphangok esetén is stabilabbnak bizonyult. Ez nem meglepő, hiszen az /a:/ ejtésének kedvez az éneklés egyik artikulációs szükségszerűsége, az állkapocsnyitás-növelés, mely az intenzív és homogén hangszínezetet fenntartó alaphangemelés eszköze magas alaphangon. Az akusztikum tekintetében ezt úgy lehet megfogalmazni, hogy az /a:/ produkcióját kevésbé befolyásolja az f_0 , hiszen az alaphang értéke az /a:/ magas első formánsát csak kifejezetten magas alaphangmagasságon éri el. Így tehát azt várhatjuk, hogy az /a:/ észlelésére is csak a kifejezetten magas alaphangú ejtés van hatással. A korábbi percepciós adatok egy része igazolja is ezt a teoretikus alapokon nyugvó elvárást (Scotto di Carlo–Germain 1985; Gottfried–Chew 1986; Sundberg 1994 egyes eredményei), ugyanakkor más kutatásokban az /a:/ percepciós instabilitásáról is beszámolnak (l. Deme 2012 összefoglaló munkáját). Az utóbb említett vizsgálathoz hasonlóan a jelen kutatásban is találkozunk az instabilitás példáival, a B4 és F5 hangmagasságok esetében ugyanis az /a:/ alacsonyabb azonosítási arányt mutatott, a tévesztések hierarchiáját legnagyobb részben az alacsonyabb első formánsú /ɒ/ uralta. Ezek a tendenciák azonban csak első ránézésre tűnnek váratlan eredménynek. A beszédhangok percepcióját vizsgáló szakirodalom ugyanis leírja, hogy a

vokálisok zártságának percepciója nem az első formáns abszolút értékében, hanem sokkal inkább annak az f_0 -hoz viszonyított tonotopikus elhelyezkedésében van kódolva: az első formáns és az f_0 frekvenciaértékének közelítése a zártabb ejtés, míg az értékek távolodása a nyitottabb ejtés érzetét kelti (Traunmüller 1981). Mindezek alapján tehát azt feltételezzük, hogy a jelen (és a korábbi) kutatás által bemutatott zártabbra tévesztések és a nyílt ejtésű hangzó perceptuális bizonytalansága az f_0 és az első formáns változó (csökkenő) viszonyából, és az ehhez kapcsolódó észlelési sajátosságból fakad.

A magánhangzó-azonosítási arányok kondíciók szerinti összehasonlításakor a mássalhangzó-környezettel összefüggésbe hozható egyetlen tendenciát is megtaláltuk: a magánhangzók ugyanis a CVC kondícióban mutatták a legszéttartóbb tendenciákat. Az /a:/ esetében itt tapasztaltuk a legmagasabb, míg az /i:/ esetében a legalacsonyabb arányokat, a két, az állkapocsnyitás szögében különböző hangzó észlelési tendenciái ebben a kondícióban tértek el a legnagyobb mértékben. Ez az eredmény egybeesik a korábbi megjegyzésünkkel, mely a magánhangzók kategorikus észlelésére hivatkozik, és paradoxnak tűnő módon, bizonyos tekintetben a hipotézisünket is alátámasztja. Az /a:/ esetében tapasztalható, a többi környezettel szemben nagyobb mértékűnek tűnő stabilitás könnyen belátható módon utalhat a mássalhangzó-környezet percepciót segítő hatására. Ugyanakkor az a tény, hogy az /i:/ már alacsonyabb alaphangon is rosszabb felismerési arányt mutat CVC kondícióban, szintén lehet éppen annak a jele, hogy a mássalhangzók segítik a hangazonosítást. Ha ugyanis az őt érintő artikulációs változások akusztikai következményeinek folyamánaképpen az /i:/ már alacsonyabb alaphangokon is egy másik percepciók kategóriába kerül, az ott tapasztalt alacsony azonosítási arányok esetleg éppen azt mutatják, hogy a hallgatók ezt a másik kategóriát ismerik fel, ráadásul jó arányban.

A produkció módja két okból is az észlelést befolyásoló tényezőnek tűnik. Egyfelől az énekelt mintákkal szemben a beszédben ejtett hangzók több esetben is (az összesített eredményekben, illetve az /u:/ esetében is) mutatták a feltételezett, de az éneklésre nem igazolódott tendenciát: a mássalhangzó-környezet és a zöngéindulás pozitív hatását. Másfelől az /a:/ és /i:/ esetében a beszélve ejtett realizációk mutatták a legmagasabb azonosítási arányokat (alaphangmagasságtól függetlenül) még azokhoz az énekelt alaphangmagasságokhoz képest is, melyek a beszéd átlagos f_0 -értékéhez közeliak voltak.

A kutatás eredményeit látva megállapíthatjuk, hogy a mássalhangzó-környezet az énekelt magánhangzók azonosítására gyakorolt hatásának vizsgálatához a jövőben nemcsak a képzéshely hatásából fakadó percepciók különbségek vizsgálata, de a mássalhangzók realizációjának és esetleges redukciójának kutatása is szükséges. Addig ugyanis nincs okunk feltételezni, hogy a mássalhangzó, illetve a koartikulációs formánsátmenetek bármilyen befolyással bírhatnak a percepciók működésekre az énekelt magánhangzók esetében, míg az sem tisztázott, hogy maga a konzonáns és a formánsátmenetek

egyáltalán megjelennek-e, és elérhetőek lesznek-e az akusztikumban, illetve a percepció számára.

Irodalom

- Boersma, Paul – Weenink, David 2013. *Praat: doing phonetics by computer [Computer program]. Version 5.3.* <http://www.praat.org>. (Letöltés ideje: 2013. júl. 9.)
- Deme, Andrea 2012. On the Hungarian sung vowels. *The Phonetician* 105–106. 73–87.
- Deme, Andrea – Grácz, Tekla Etelka – Jankovics, Julianna 2013. Obstruent voicing in singing. (Előadás) *15th International Summer School of Psycholinguistics*, 2013. május 26–30; Balatonalmádi, Hungary.
- Evgrafova, Karina – Evdokimova, Vera 2012. Perception of Russian vowels in singing. *Baltic HLT Frontiers in Artificial Intelligence and Applications*. IOS Press 247. 42–49.
- Garnier, Maëva – Henrich, Natalie – Smith, John – Wolfe, Joe 2010. Vocal tract adjustments in the high soprano range. *Journal of the Acoustical Society of America* 127/6. 3771–3780.
- Goldstein, Ursula 1980. *An articulatory model for the vocal tracts of growing children*. D.Sc. thesis. Massachusetts Institute of Technology, Cambridge, MA.
- Gottfried, Terry L. – Chew, Stephen L. 1986. Intelligibility of vowels sung by a countertenor. *Journal of the Acoustical Society of America* 79/1. 124–30.
- Hollien, Harry – Mendes-Schwartz, Ana P. – Nielsen, Kenneth 2000. Perceptual confusions of high-pitched sung vowels. *Journal of Voice* 14/2. 287–298.
- R Core Team 2013. *R: A Language and Environment for Statistical Computing*. <http://www.R-project.org>.
- Ryalls, John. H. – Lieberman, Philip 1982. Fundamental frequency and vowel perception. *Journal of the Acoustical Society of America* 72/5. 1631–1634.
- Scotto di Carlo, Nicole – Germain, Aline 1985. A perceptual study of the influence of pitch on the intelligibility of sung vowels. *Phonetica* 42/2. 188–97.
- Sjölander, Kåre – Beskov, Jonas 2009. *Wavesurfer*. <http://www.speech.kth.se/wavesurfer>.
- Smith, Lloyd A. – Scott, Brian L. 1980. Increasing the intelligibility of sung vowels. *Journal of the Acoustical Society of America* 67/5. 1795–1797.
- Strange, Winifred – Verbrugge, Robert R. 1976. Consonant environment specifies vowel identity. *Journal of the Acoustical Society of America* 60/1. 213–224.
- Sundberg, Johan 1987. *The science of the singing voice*. Northern Illinois University Press, Illinois.
- Sundberg, Johan 1994. Perceptual aspects of singing. *Journal of Voice* 8/2. 106–122.
- Traunmüller, Hartmut 1981. Perceptual dimension of openness in vowels. *Journal of the Acoustical Society of America* 69/5. 1465–1475.

A kutatás alatt a szerző a Balassi Intézet Campus Hungary programjának ösztöndíjasa volt. A felvételek a Kempelen Farkas Beszédkutató Laboratóriumban készültek (MTA Nyelvtudományi Intézet, Budapest), az adatokat részben a Department of Speech, Music and Hearing (KTH Royal Institute of Technology, Stockholm) intézményben elemeztük, professzor Johan Sundberg (KTH Royal Institute of Technology) szíves segítségével.

EGYSZERRE BESZÉLÉSEK DETEKTÁLÁSA A BESZÉLŐDETEKTÁLÁS JAVÍTÁSÁHOZ

Beke András

Bevezetés

A társalgás a spontánbeszéd-technológia speciális esete, mivel a gépi beszédfelismerő rendszerek számára nehezebb az olyan beszéd típusok dekódolása, ahol több beszélő társalog egymással. Ezért megnőtt az igény a gépi beszélődetektálásra is. A társalgás során a monologikus beszédre jellemző akusztikai és nyelvtani szabályok nagyszámú varianciája mellett újabb nehézségek jelennek meg. Ezek lehetnek a társalgást jellemző egységek, mint például a beszédforduló, az egyszerre beszélés, a nonverbális jelek (nevetés), ezért a beszélődetektáláskor valamennyiük modellezésére szükség van (Boakye et al. 2008, 2011; Zelenák et al. 2010).

A társalgás alapegysége a beszédforduló (angol terminusban *turn*). A beszédforduló során a társalgás egyik résztvevője beszél, amíg át nem adja, vagy amíg át nem veszik tőle a beszéd jogát (Sacks et al. 1974). A társalgásban az elméletek szerint egyszerre csak egy beszélő beszél, ezért az átfedések, vagyis az egyszerre beszélések és a hosszabb szünetek hibának minősülnek, amelyeket a beszélők igyekeznek javítani (Boronkai 2008a, b). Heldner és Edlund (2010) tanulmányában a beszélőváltások lehetséges módozataival foglalkozott akusztikai aspektusból. Az elméletek szerint három lehetséges módon történhet beszélőváltás: a) szünet van a két beszélő megnyilatkozása között, b) a két beszélőtől származó megnyilatkozás időben átfedi egymást, vagy c) sem szünet, sem átfedő beszéd nem történik. Sacks, Schegloff és Jefferson (1974) munkájukban megkülönböztetnek időközt (hosszabb vagy megnyúlt szünet), hiányosságot (rövidebb szünet) és átfedéseket (átfedő beszéd), valamint nincs-szünet-nincs-átfedést.

Az átfedő beszéd aránya a spontán társalgásokban meglehetősen nagyinak mondható (Grácsi–Bata 2010). Beattie (1983) a beszélőváltásokat elemezve kimutatta, hogy a két résztvevős angol társalgásban 11%-ban fordul elő egyszerre beszélés (azaz a beszédpartner közbevág), több beszélőnél ez az arány már 31%. Az újabb kutatások ezeket az arányokat igazolták. Cetin és Shriberg (2006) angol korpuszokat vizsgálva adatolta, hogy az átfedő beszéd átlagosan 10–13%-át teszi ki a társalgásoknak. A hazai kutatásokban Markó (2005) 6%-ot állapít meg a teljes beszéd és az átfedő beszéd arányaként négybeszélős spontán társalgásban. Bata (2009) 1,7–3%-ot adatolt kutatásában, ahol spontán társalgásokat elemzett. Ez a magas előfordulási szám az át-

fedő beszéd funkciójából adódik. A társalgás során ugyanis az egyszerre beszélés kettős funkciót tölt be. Egyrészt megerősítő szerepe van (pl. *igen, aha, ühüm*, amelyek háttércsatorna-jelzések), másrészt versengő funkciójú, amikor a társalgás egyik szereplője át kívánja venni a szót, és már az alatt elkezdí a beszédét, amikor az aktuálisan beszélő még nem fejezte be mondanivalóját (Iványi 2001; Hámosi 2006; Bata 2009).

A társalgásban a beszélők tapasztalati úton érzékelik a beszélőváltásra alkalmas helyeket, amelyekhez bizonyos szuprasegmentális, szemantikai, pragmatikai kulcsok kombinációját alkalmazzák/alkalmazhatják. A beszélőváltások vizsgálatára számos kutatás történt. Vizsgálták a beszélőváltások dallammenetét (Wells–Peppé 1996), előre jelezhetőségét (De Rulter et al. 2006), a zöngé minőségét (Ogden 2004), a magyar nyelvben a fonetikai megvalósulásukat és szintaktikai szerkezetüket (Markó 2005; Bata 2009; Grácsi–Bata 2010).

A diskurzuselemzés felől jelentős mennyiségű munka foglalkozott már az egyszerre beszélésekkel (vö. Cetin–Shriberg 2006). Az átfedő beszéd több szempontból is jelentős. A diskurzuselemzésben fontos kérdés, hogy mikor következik be az egyszerre beszélés a társalgó felek szociális viszonyaitól, ismertségi fokától és egyéb tényezőktől függően, és hogy ezek az átfedő beszédek milyen szintaktikai, pragmatikai, illetve fonetikai formában jelennek meg. Fontos szerepük van továbbá a spontán beszéd automatikus felismerésében is, hiszen az egyszerre beszélések a gépi beszédfelismerés számára korlátozottan – illetve lényegében egyáltalán nem – feldolgozható szakaszai a beszédnek (Boakye et al. 2008). A beszélődetektálásban a beszélői modell kialakítása során az átfedő beszédrészek mint zaj jelentkeznek. Ez azért lehetséges, mivel az átfedő részekben nem csak egy beszélő jelenik meg akusztikailag, ami az egyes beszélői modell egységességét gyengítheti, csökkentve ezzel a végleges beszélődetektálási eredményt. Ezért elengedhetetlen, hogy az átfedő részek gépi úton automatikusan azonosíthatók legyenek.

A beszélődetektálásban kimutatták, hogy a legtöbb hiba szignifikánsan azon részekben történik a felvételekben, ahol egyszerre beszélés található. Wooters és Huijbert (2007) munkájukban azt írták le, hogy a beszélődetektálás hiba arányának 17%-át a téves elutasítások száma adja, amelyet az átfedő beszédrészek okoznak.

Az egyszerre beszéléseket modellező munkák száma relatíve kevés, és azok közül is csak néhány kutatásban mutatták ki, hogy csökkenti a beszélődetektálási hibaarányát (DER: Diarization Error Rate) (Boakye et al. 2008; Boakye 2008; Trueba–Hornero 2008).

Moattar és Homayounpour (2006) a társalgásban megjelenő egyszerre beszélést a hang periodicitásából ítélték meg. A vizsgálat során azt figyelték meg, hogy ahol a beszéd nem mutatott periodicitást a Fourier-spektrumban, ott jelent meg az egyszerre beszélés. Boakye és munkatársai (2008) kimutatták, hogy az átfedő beszédet MFCC és más akusztikai paraméterekkel

GMM/HMM-mel modellezve 7,4%-ban csökkenteni lehetett a detektálási hiba arányát a beszélőazonosításban. Ugyancsak Boakye és munkatársai (2011) amerikai angol spontán társalgási korpuszban vizsgálták az átfedő beszédresek automatikus osztályozhatóságát a beszélődetektáló rendszerek javítása érdekében. Akusztikai jellemzőként MFCC-t, RMS-energiát, LPC-analízist és még számos más, a zöngeminőségét jellemző eljárást alkalmaztak. Ezeket dimenziócsökkentették, és GMM-mel mintaillesztették. A hasonlóság méréséhez Kullback–Leibler-távolságot számoltak. Ezzel az eljárással kimutatták, hogy szignifikánsan csökkenthető a tévesztési arány a beszélődetektálás során a spontán társalgásokban. Otterson és Ostendorf (2007) munkájukban elméleti megközelítésben kimutatták, hogy az átfedő beszéd osztályozásával javítani lehet a beszélődetektálás eredményét. Az általuk létrehozott osztályozót azonban nem tesztelték beszélődetektálóban. Trueba-Hornero (2008) munkájában már egy valós átfedőbeszéd-detektálót hozott létre, és tesztelt beszélődetektálóban. A legtöbb munka azonban nagyon magas hibaértékekről számol be, ami mutatja a feladat nehézségét (Boakye et al. 2008; Boakye 2008). Ezen alkalmazások HMM-GMM-et használnak, amelyben három modellt hoznak létre: nem beszéd, nem átfedő beszéd és átfedő beszéd. Az eredmények azt mutatták, hogy a legjobb eredményük alapján a pontosság (precision) 58%, míg a fedés (recall) 19% volt. Az alacsony pontossági és fedési értékek mellett is 10%-os relatív DER-csökkenést tudtak elérni az árfedő beszédresek detektálásával. Jóllehet ezek az eredmények messze elmaradnak a várttól, becslések szerint azonban az ideális egyszerre beszéléseket detektáló algoritmussal a DER 37%-kal lenne csökkenthető, ezért ezen a területen még igen sok fejlesztésre van szükség (Wooters–Huijbert 2007).

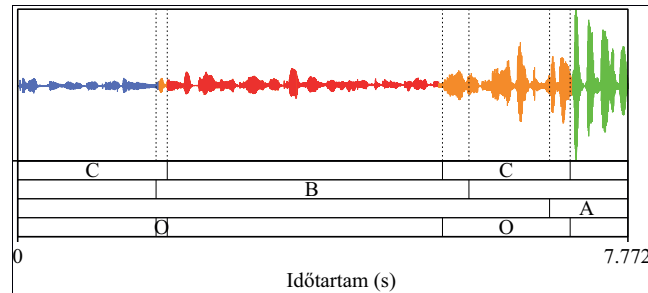
A jelen kutatás célja, hogy a spontán társalgásokban modellezze az egyszerre beszéléseket, és automatikus osztályozó algoritmussal különítse el azoktól a beszédszakaszoktól, ahol csak egy társalgó beszél. A kutatás további célja, hogy az egyszerre beszélések detektálásával javítsuk a beszélődetektálás eredményeit. Hipotézisünk szerint az átfedő beszéd jellegzetes akusztikai szerkezettel rendelkezik, ezért létrehozható egy automatikus osztályozó algoritmus. Ugyanakkor feltételezzük, hogy a háttérszóra-jelzések okozzák majd a legtöbb hibát az osztályozáskor. Feltételeztük továbbá azt is, hogy az egyszerre beszéléseket detektáló algoritmussal a beszélődetektálóba való implementálásával a DER értéke csökkenthető.

Anyag, kísérleti személyek, módszer

A vizsgálatban a BEA adatbázisból (Gósy 2012) 100 társalgást választottunk ki, amely 55 órányi társalgást jelent. A társalgásokban minden esetben három személy vett részt. Ebből két társalgó állandó volt (2 nő, életkoruk 33 év). A harmadik személy 43 férfi és 57 nő közül került ki, átlagos életkoruk 35 év.

A felvétel minősége laboratóriumi körülményekhez hasonló. A felvételt egy AT-4040 irányított mikrofonnal, egy csatornára rögzítették 44 kHz-en, amelyet újrámintavételeztünk 16 kHz-en. A BEA korpusz alapvető céljának megfelelően az adatközlőhöz volt legközelebb a mikrofon, így az ő beszédjele volt a legerősebb, míg a kísérletvezető, illetve egy másik bevont személy beszédjele gyengébb volt. Ez megnehezítette az egyszerre beszélések automatikus osztályozását. Lehetőség lett volna normalizációs eljárásokat használni, de feltehetően a zajt is felerősítette volna.

A társalgásokban manuálisan jelöltük azokat a részeket, ahol egyszerre több adatközlő beszél, illetve azokat a részeket, ahol csak egy beszélő beszél (1. ábra).



1. ábra

Az átfedő beszéd illusztrálása (A, B, C = beszélők, O = egyszerre beszélés)

A 100 beszélő spontán társalgásaiban összesen 8056 olyan időintervallum található, ahol kettő vagy annál több résztvevő szólal meg egyszerre, vagyis ahol átfedő beszéd van. Ezen intervallumok összhossza közel 7 óra, a teljes korpusz 12%-a.

Módszer

Jóllehet az egyszerre beszélések automatikus osztályozása egyszerű feladatnak tűnik, megvalósítása korántsem triviális. A beszélődetektálás egyik alapfeladata, mégis csak néhány olyan tanulmány ismert, amely megfelelő eredménnyel tudta megvalósítani az egyszerre beszélések automatikus osztályozását (vö. Boakye et al. 2008).

A jelen kutatásban egy ANN/SVM hibrid rendszert (Artificial Neural Network/Support Vector Machine: Mesterséges Neuronháló/Szupportvektor-gép) hoztunk létre az egyszerre beszélések automatikus osztályozásához.

Az osztályozás során az első lépés a lényegkiemelés, amelynek fő feladata, hogy a beszédjelből olyan információkat vonjuk ki, amellyel jól megragadhatók az egyszerre beszélések. Mivel nem ismert, hogy mely akusztikai paraméter mentén különülnek el az átfedő beszédrészek és a nem átfedő beszédrészek, több akusztikai jellemzőt is teszteltünk, mint az FFT-spektrum, rész-

sáv-energia (subband-energy), MFCC és Mel-skála szerinti logaritmikus szűrőbank. A jellemzők jobb reprezentálásához főkomponens-analízist (PCA: Principal Component Analysis) használtunk.

Az osztályozás második fontos lépése a mintaillesztés, amelyben két fontos részfeladatot kell megoldani: (i) osztályozás, vagyis melyik beszédrészlet-modell a legvalószínűbb az adott időpillanatban; (ii) időillesztés: melyik időszegmenst rendeljük az egyik vagy a másik modellhez. Ennek megvalósításához a beérkező mintát, vagyis vektorsorozatot (statisztikai úton becsült) valószínűségmodell-struktúrához illesztjük. Az akusztikus modell létrehozásához legtöbbször a Gauss-keverék modellt (GMM: Gaussian Mixture Model) használják. Bár az akusztikus modell létrehozásában igen széles körben és kiválóan alkalmazható, mégis számos hátránya létezik. Az egyik hátránya, hogy előzetes feltételeknek kell megfelelniük az adatoknak a becslést megelőzően – ilyen követelmény a normál eloszlás. A GMM alternatívájaként léteznek más mesterséges neuronhálók, mint a MLP (Multilayer Perceptron; Bourlard–Morgan 1993). Az elmúlt években az ANN egy új fajtája jelent meg: ún. mély neuronhálók, amely vizsgálatok szerint igen jól alkalmazhatók a beszédhang-felismerésben (Dahl et al. 2010; Tóth–Grósz 2013). A mély neuronhálók elsősorban abban különböznek az előző neuronhálóktól, hogy általában nem egy, hanem 3–9 rejtett réteget használnak. A több rejtett réteg tanításához újfajta tanulóalgoritmust is fejlesztettek. A jelen kutatásban a mély neuronhálókat az akusztikai jellemzők előfeldolgozásához használtuk. A tényleges osztályozást LS-SVM-el végeztük el, amely az SVM egyik változata. Korábbi tanulmányok kimutatták, hogy az ANN és az SVM algoritmusok kombinációi jól alkalmazhatók automatikus osztályozásához (Bellili et al. 2001).

1. Jellemzőkinyerés. Az egyszerre beszélések jó megfeleltethetőségéhez az akusztikai beszédjelből különböző jellemzőket nyertünk ki. A mély neuronhálókat a hangfelismerésben oly módon szokás alkalmazni, hogy előfeldolgozásként a hangot valamilyen képformátumúvá alakítjuk. Ennek egyik legegyszerűbb formája a különféle spektrumok, vagy az emberi hangok esetén a Mel-skálázott spektrogram. A jelen kutatásban négy különböző eljárást teszteltünk a beszédhang képpé alakítására.

(i) A **spektrum (SP)** kiszámolásához 256 pontos FFT-analízist használtunk Hamming-ablakkal, az ablak hossza 32 ms volt (8000 Hz-es mintavételezés esetén), amelyet 10 ms-onként léptettünk. A jellemzővektor hossza ebben az esetben 257 minden egyes 10 ms-os időkeretre. Mivel a 257 dimenzió igen nagy, ezért PCA-val (Principal Component Analysis: főkomponens-analízis) lecsökkentettük 80-ra.

(ii) A **Mel-frekvenciás kepsztrális (MFC)** együtthatók kinyeréséhez a PLP-RASTA csomagban található, Matlab szoftverkörnyezetre írt MFCC-algoritmust használtuk (vö. Ellis 2005). A jellemzők száma egy-egy időkeretben 39: a szokásos 12 MFCC koefficiens + az energia logaritmusa + ezek el-

ső két deriváltja ($13 \cdot 2 = 26$). Ezt a 39 paramétert 10 ms-onként 25 ms-os, 50%-ban átlapolódó időkeretekben kimértük. A jellemzővektor hossza így 39 minden egyes 10 ms-os időkeretre.

(iii) A **Mel-skála szerinti logaritmikus szűrőbank** (MSL) számítása ugyanúgy történik, ahogyan az MFC kiszámítása. A különbség abban áll, hogy a Mel-frekvenciás szűrés után vesszük annak logaritmusát, de nem végezzük el a kepsztrális transzformációt. Ennek kiszámítása szintén 12 koefficiens + az energia logaritmusa + ezek első két deriváltja ($13 \cdot 2 = 26$). Ezt a 39 paramétert 10 ms-onként 25 ms-os, 50%-ban átlapolódó időkeretekben kimértük. A jellemzővektor hossza így 39 minden egyes 10 ms-os időkeretre.

(iv) A **részsáv-energiát** (RSE) úgy számoltuk ki, hogy a spektrumot 20 részsávra bontottuk, majd mind a 20 részsávban kiszámoltuk a jel energiáját. A folyamat végén a 20 elemű vektort DCT-vel (Discrete Cosine Transformation) dimenziócsökkentettük 12-re (vö. Sarikaya et al. 1998).

Mindegyik jellemző esetén a különféle zajok – elsősorban a konvolúciós zajok (pl. csatornatorzítás) – hatását mérséklendő további transzformációs lépést alkalmaztunk: kepsztrális átlagkivonást (CMS: cepstral mean subtraction).

Mivel a következő lépésben neurális hálózatot alkalmazunk, ezért az adatokat 0 és 1 közé normalizáltuk.

2. Lényegkiemelés

Korlátozott Boltzmann-gép. Az elmúlt években számos kísérlet bizonyította, hogy a gépi látásos módszerek jó eredménnyel alkalmazhatók beszéddel kapcsolatos problémák megoldására (Dahl et al. 2010). A gépi látásos módszerek egyik legtöbbet használt algoritmus a Konvolúciós Hálózatok. A Konvolúciós Hálózatok hierarchiát alkotva több szintből épülnek fel, ahol az alsóbb szinteken csak egy kis részét látják a képnek, erről a részletről lokális jellemzőket nyernek ki, amelyet a felsőbb szinteknek továbbítanak, és egyre feljebb jutva az egyes szinteken egyre általánosabb jellemzőket állapítanak meg.

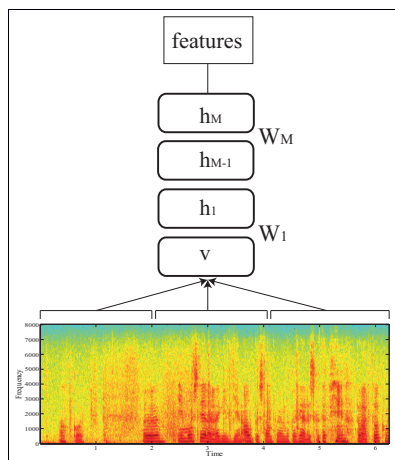
A korlátozott Boltzmann-gép (RBM: Restricted Boltzmann Machine) alapvetően két különböző réteget tartalmaz: látható és rejtett réteget. A korlátos jelző arra utal, hogy a neuronok között csak akkor van összeköttetés, ha az egyik a látható, a másik pedig a rejtett réteghez tartozik. Az azonos rétegbe tartozó neuronok között nincs összeköttetés.

A súlyok az egyes kapcsolatok között, illetve a neuronokhoz tartozó eltolásértékek (biasok) egy véletlen eloszlást definiálnak a látható réteg neuronjainak állapotait tartalmazó vektorok felett, amelyet egy energiafüggvény segítségével írhatunk le. Az alap energiafüggvény bináris adatok eloszlásának leírására alkalmas. Mivel a jelen kutatásban az RBM bemeneti vektorai valós értékűek, ezért az RBM-eknek a Gauss-Bernoulli RBM változatát használjuk.

A korlátozott Boltzmann-gép tanító algoritmus a CD-algoritmus (kontrasztív divergencia). A CD-algoritmus felügyelet nélküli tanulást végez,

amely a „maximum likelihood”-tanítás közelítését adja. Ezt a folyamatot az RBM előtanításának nevezzük (Tóth–Grósz 2013).

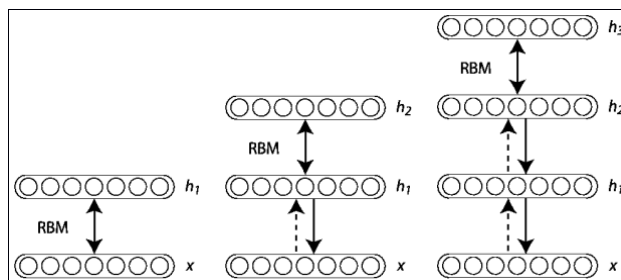
A jellemzők kinyerése után korlátozott Boltzmann-géppel emeltük ki a lényegét az akusztikai jellemzőkből. A korlátozott Boltzmann-gépet szokás jellemzőkinyerésre is alkalmazni – főként a képfeldolgozásban –, amely ebben az esetben nemellenőrzött tanulási eljárással működik (2. ábra).



2. ábra

Jellemzőkinyerés korlátozott Boltzmann-géppel

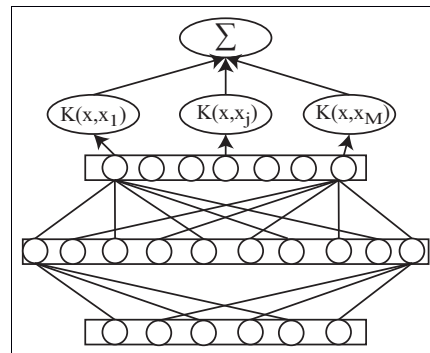
A korlátozott Boltzmann-gép igen jónak bizonyult a képi feldolgozásban. Az RBM előnye, hogy könnyedén mély neuronná lehet alakítani, ha az egyes RBM-eket összekötjük, előállítva ezzel egy hierarchikus tanulási láncot, így segítve a magasabb szintű struktúrák kinyerését az adatokból (3 ábra).



3. ábra

Korlátozott Boltzmann-gép és a belőle felépített mély neuronháló
(Deep Neural Network)

Az RBF tanítása után a rejtett rétegek aktivációs értékeit használtuk fel az átfedő beszédrészek és nem átfedő beszédrészek automatikus osztályozásához, amelyet szupportvektorgéppel valósítottunk meg (4. ábra).



4. ábra
Szupportvektorgép mély neuronhálózval előtanítva

Az RBM előtanítási paraméterei. Az RBM előtanításához az akusztikai paramétereket 15 keret hosszúságú csúszóablakkal nyertük ki. Mindegyik összefüggő ablakot felhasználtuk az RBM tanításához. Az RBM látható egységeinek száma a jellemzővektor dimenziószámának a keret hosszával képzett szorzata. Minden egyes audio szegmensre az érvényes konvolúcióval kifejezve $m - n + 1$ összefüggő ablak adódik, ahol m a keretek száma, n a csúszóablak hossza. A mély rétegű neurális hálózatok (DBN: Deep Belief Network) létrehozásához 1–3 RBM-et kapcsoltunk össze úgy, hogy a megelőző rejtett réteg aktivációja a következő látható réteg bemenete.

Az első RBM-ben (H1) a unitok száma 300.

A második RBM-ben (H2) a unitok száma 600.

A harmadik RBM-ben (H3) a unitok számát 300–900-ig növeltük 100 unitonként.

Minden egyes rétegben energiafüggvényként a Gauss-Bernoulli algoritmust használtuk. A batch mérete 100 volt, amely a kötegelt tanítás mérete. Az első rétegben 50 iterációt használtunk, a többi rétegben 25-öt.

Az RBM megvalósításához Kyung Hyun Cho a Matlab-ban írt GitHub toolbox-át használtuk (Cho 2014).

3. Osztályozás

Az átfedő és nem átfedő beszédrészeket szupportvektorgéppel (SVM) kernelfüggvényként radiális bázisfüggvényt (RBF) alkalmazva osztályoztuk.

Szupportvektorgép (Support Vector Machine). Az SVM olyan matematikai konstrukció, amelyet döntési problémák megoldásához szoktak alkalmazni. Alapverziója a lineáris osztályozók családjába tartozik, de bináris osz-

tályozási problémák megoldására alkalmas. A többi lineáris osztályozóhoz képest az a fő ismérve, hogy nemcsak egyszerűen olyan hipersíkot (más néven vágási síkot) keres, amely elválasztja a pozitív és a negatív tanítómintákat, hanem ezek közül a legjobbat kutatja, vagyis intuitíve azt, amelyik a két osztály mintái között éppen „középen” fekszik (Borges 1998). Az SVM-et alapvetően lineárisan szeparálható esetekre találták ki. A valóságban azonban a legtöbb probléma nemlinearitása olyan nagyságrendű, hogy az osztályozó nem lesz hatékony.

Ennek a problémának a megoldására az adatokat nagyobb dimenziójú térbe transzformáljuk, ahol az adathalmaz már lineárisan szeparálható. Az erre képes matematikai függvényeket kernel- vagy magfüggvényeknek nevezzük. A gyakorlatban a következő magfüggvényeket szokták alkalmazni: polinominális, radiális bázisfüggvény, kétrétegű perceptron.

A jelen kutatásban az SVM egy változatát használtuk, amely az LS-SVM (Least Square Support Vector Machine, Suykens et al. 2002). Ez a típus abban tér el az alap SVM-től, hogy az idő- és energiaigényes kvadratikusan programozás helyett lineáris egyenletrendszerre vezeti vissza a megoldandó problémát, ezáltal a számítási idő jelentősen csökken.

A kész osztályozó kiértékeléséhez a tesztadalmat használhatjuk. Vizsgálatunkban az osztályozáshoz az LS-SVM függvénykészletet használtuk (Matlab implementáció, Chih-Chung–Chih-Jen 2012) az úgynevezett radiális bázis (RBF – Radial Basis Function) kernelfüggvénnyel. Így a supportvektorgépnek két szabadon állítható paramétere van: C a hibázási paraméter (penalty parameter) és γ az RBF kernelfüggvény (Gauss-függvény) szórásparamétere. Érdekes először egy úgynevezett keresztvalidációs eljárással (cross-validation) és egy optimalizáló eljárással (simplex method) kizárólag a tanítóhalmazon beállítani az SVM-tanítás említett paramétereit (Hsu et al. 2003). A fentiek elvégezve az SVM számos lehetséges C és γ paraméterpárjára (kimerítő keresés, grid-search) megtalálhatjuk az optimális beállítást, vagyis amikor az SVM a legnagyobb felismerési arányokat éri el. Hsu, Chang és Lin (2003) szerint a C és γ értékeket az alábbi tartományokban érdemes keresni:

$$C: \{2^{-5}; 2^{-3}; \dots; 2^{13}; 2^{15}\}$$

$$\gamma: \{2^{-15}; 2^{-13}; \dots; 2^1; 2^3\}$$

Az SVM tanítási paramétereit. Az átfedő és nem átfedő beszédrészek osztályozásához a korpusz minden beszédszegmensére kinyerjük az akusztikai jellemzőket, majd a tanítóhalmaz értékeivel tanítjuk be az osztályozót.

Az SVM tanításához a 8056 átfedő beszédszegmens 2/3-át, vagyis 5370-et használtunk fel, míg a teszteléshez az 1/3-át, amely 2386 szegmenst jelent. A korpuszban az átfedő beszédszegmensek előfordulása alacsonyabb volt, ezért a nem átfedő beszédrészek számát ehhez igazítottuk a tanító adatbázisban

(random kiválasztási módszerrel). Erre azért volt szükség, hogy az algoritmus ne tanuljon rá jobban az egyik csoportra.

Ahhoz, hogy az SVM-et alkalmazni tudjuk, először az adatokat azonos dimenziójúra kell hoznunk. Mivel nem minden audioszegmens ugyanolyan hosszúságú, ezért a bemenő jellemzővektorok dimenziója nem egyenlő. Ennek kiküszöbölésére az egyes audioszegmensek kereteire statisztikai jellemzőket számolunk (átlag és szórás).

Az SVM bemeneti vektora tehát (i) a spektrumra: 2×80 ; (ii) az MFCC-re 2×39 ; (iii) részsáv-energiára 2×12 .

Az SVM RBF-függvényének két szabad paraméterét, a C -t és a γ -t háromszoros keresztvalidációval és softmax függvénnyel optimalizáltuk.

Az osztályozás kiértékelése

(i) DET (Detection Error Tradeoff). Az osztályozásra alkalmazott algoritmusok működésének kiértékelésére és összehasonlítására a DET (Detection Error Tradeoff, Martin et al. 1997) kiértékelő algoritmust használtuk. A DET kiértékeléséhez először bemutatjuk a bináris osztályozás esetén a tévesztési mátrixot (1. táblázat).

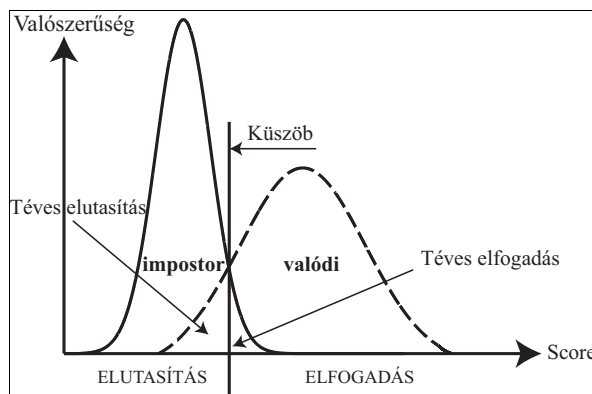
1. táblázat: A bináris osztályozás esetén a tévesztési mátrix

		Aktuális feltétel	
		Pozitív	Negatív
A teszt eredménye	Pozitív	A feltétel teljesül + pozitív teszt = TP (True Positives)	A feltétel nem teljesül + pozitív teszt = FP (False Positives)
	Negatív	A feltétel teljesül + negatív teszt = FN (False Negatives)	A feltétel nem teljesül + negatív teszt = TN (True Negatives)

A bináris osztályozáskor megkülönböztetünk első- és másodfajú hibát. Az elsőfajú hiba a téves elfogadás (False Acceptance Rate: FAR; False Positives). A jelen munka során a téves elfogadásról akkor beszélünk, ha a beérkező szegmens nem átfedő beszéd, de annak fogadja el a gép. A másodfajú hiba a téves elutasítás (False Rejection Rate: FRR; False Negatives) (5. ábra). A jelen munka során a téves elutasításról akkor beszélünk, ha a beérkező szegmens átfedő beszéd, de nem fogadja el annak a gép.

Az osztályozó egy-egy összehasonlítás során a hangmodelleket összeveti az aktuális jellemzőkkel, és mintánként egy hasonlósági számot képez (score), aztán sorba állítja az eredményt a csökkenő score szerint, és döntést hoz, hogy az első helyen levő találat-e vagy sem. A küszöbérték (threshold) alapján dönt a találatról: ha az első „score” érték alacsonyabb a küszöbértéknél, akkor nincs találat (NOHIT), ha magasabb, akkor van találat (HIT). Ekkor felmerül az a kérdés, hogy milyen küszöbértéket állítsunk be, hogy az osztályozás a lehető legjobb legyen. Ennek megoldására léteznek különböző

technikák, mint a ROC (Receiver Operating Characteristic), vagy a DET (Detection Error Tradeoff). A DET-ben úgy választjuk meg a küszöbértéket, hogy az elsőfajú hiba és a másodfajú hiba egyenlő legyen. Ezt úgy hívják, hogy Equal Error Rate.

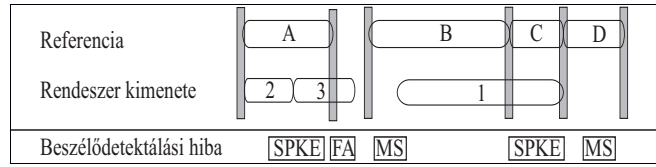


5. ábra

A bináris osztályozáskor fellépő hibák sematikus ábrázolása

(ii) **Beszélődetektálási hibaarány (DER = diarization error rate).** A beszélődetektálás kiértékeléséhez a NIST munkatársai által fejlesztett DER-algoritmust használtuk, amelyet a NIST az RT kiértékelésekor alkalmazott (NIST Fall Rich Transcription 2006). A DER-t tulajdonképpen úgy értelmezzük, mint azt a törési időt, amely nem tulajdonítható helyesen sem a beszélőnek, sem a nem beszélőnek. Ennek mérésére az MD-eval-v12.pl-t, a NIST MD-eval-v12 DER kiértékelő szkriptjét használtuk.

Mivel a váltási pontok meghatározása a feladat, a rendszer hipotéziseként a beszélődetektálás kimenetében nem kell explicit meghatározni a beszélő nevét vagy identitását, ezért a beszélőkhöz rendelt azonosító címkéknek nem kell azonosnak lenniük a bemeneti (kézi) címkében és a kimeneti (automatikus) címkében. Ez a feladat tehát nem olyan, mint a beszéd-nem beszéd automatikus címkézése, amely során a szegmenst azonosító címkének egyezni kell a bementi és a kimeneti címkében (6. ábra).



6. ábra

A DER kiértékelési módszer sematikus ábrázolása

A kiértékelő szkript először megtalálja az optimális, egy-az-egyben átfedést az összes beszélői címke azonosítóira a referencia- és az automatikus címke között. Ez teszi lehetővé az egyezés mérését a különböző azonosítóval rendelkező két címkesor között. A DER értékét a következőképpen számoljuk:

$$DER = \frac{\sum_{s=1}^S dur(s) \cdot (\max(N_{ref}(s), N_{hyp}(s)) - N_{correct}(s))}{\sum_{s=1}^S dur(s) \cdot N_{ref}}$$

ahol az S az összes beszélői szegmens száma, ahol mind a hipotetikus, mind a referencia címke tartalmazza ugyanazokat a beszélői párokat. Ezt úgy kapjuk meg, hogy összehasonlítjuk a hipotetikus, illetve a referencia-beszédfordulókat. A N_{ref} és a N_{sys} kifejezések a beszélők számát jelölik a beszédsgemensekben, az $N_{correct}$ a beszélők számát mutatja, amely a helyes találatokat jelenti a referencia- és a hipotetikus címkesor között. A címkesorban a nembeszéd-részeket 0 beszélőnek jelölik. Ha mind a beszéd-, mind a nembeszéd-szegmensek helyesen lettek azonosítva, akkor a hiba értéke 0. A DER-hiba tulajdonképpen különböző módon létrejött hibák összege:

1) Beszélőhiba: a helytelenül azonosított beszélői azonosítók a teljes időtartam arányában. Ez a típusú hiba nem veszi figyelembe a beszélők átfedését, vagy bármilyen más hibát, ami a nembeszéd-részek azonosításából fakad. Ezt a következőképpen írhatjuk fel:

$$E_{spkr} = \frac{\sum_{s=1}^S dur(s) \cdot (\min(N_{ref}(s), N_{hyp}(s)) - N_{correct}(s))}{T_{score}},$$

ahol a T_{score} teljes időtartama a kiértékeléshez használt fájloknak.

2) Téves riasztások száma: teljes időtartamra vetítve a referenciacímkében a nem beszéd szerepel, de az automatikus címkesorban beszélőnek azonosított a szegmens. A következőképpen írhatjuk fel:

$$E_{FA} = \frac{\sum_{s=1}^S dur(s) \cdot (N_{ref}(s) - N_{hyp}(s))}{T_{score}} \forall (N_{hyp}(s) - N_{ref}(s)) > 0,$$

amit csak azon szegmensekben mérünk, amely a referenciacímkeben nem-beszéd részeként szerepel.

3) Téves elutasítások száma: a teljes időtartamra vetítve a referenciacímkeben a beszélő szerepel, de az automatikus címkesorban nem beszédnek azonosított a szegmens. A következőképpen írhatjuk fel:

$$E_{MISS} = \frac{\sum_{s=1}^S dur(s) \cdot (N_{ref}(s) - N_{hyp}(s))}{T_{score}} \forall (N_{ref}(s) - N_{hyp}(s)) > 0,$$

amit csak azon szegmensekben mérünk, amely a hipotetikus címkében nembeszéd-részeként szerepel.

4) Egyszerre beszélések: a teljes időtartamra vetítve, amikor több beszélő beszél egy szegmensben, amely nem tartozik egy beszélőhöz sem. Ez a fajta hiba általában az E_{MISS} -hez vagy az E_{FA} -hoz tartozik. Ez a hiba függ attól, hogy a referencia- vagy a hipotetikus címkesorban szerepel-e az egyszerre beszélés. Ha mindkettőben, akkor E_{spkr} -hez tartozik.

Felírva az összes lehetséges hibát, a DER a következőképpen áll össze:

$$DER = E_{spkr} + E_{MISS} + E_{FA} + E_{ovl}$$

Amikor a kiértékelést végezzük, egy olyan időbeli határsávot használunk minden referenciában lévő beszédfordulóra, amely bizonyos pontatlanságot enged meg az automatikus címkézésnek. A NIST ezt az időbeli határsávot ± 250 ms-ban határozta meg. A NIST DER kiértékelő script megadja minden egyes referencia-hipotetikus szegmentációra a DER értékét, illetve az összes kiértékeléshez használt fájlra ad egy súlyozott átlagot.

Eredmények

Az egyszerre beszélések időtartama 12%-át teszi ki a teljes korpusznak, míg a szünetek időtartama 10,9%-át; a beszédrészek tehát a teljes korpusz 77,1%-a. Az átlagos átfedőbeszéd-arány a felvételekben 21,84%.

A jelen kutatásban teszteltük, hogy a négy akusztikai paraméter közül melyikkel lehet elérni a legjobb eredményt. Továbbá teszteltük azt is, hogy hogyan változik az eredményünk annak függvényében, hogy a mély rétegű neuronhálózat harmadik rétegében hány neuront használunk.

Az eredmények azt mutatják (2. táblázat), hogy a négy akusztikai paraméter [FFT spektrum (SP); Mel-frekvenciás kepsztrális (MFC) együtthatók; Mel-skála szerinti logaritmikus szűrőbank (MSL); részsáv-energia (RSE)]

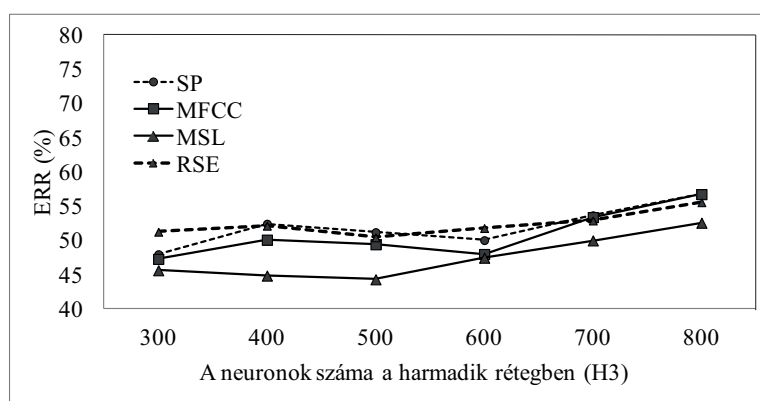
közül a legjobb teljesítményt akkor kaptuk, ha jellemzőként a Mel-skála szerinti logaritmikus szűrőbankot alkalmaztuk. Ekkor az Equal Error Rate (EER) átlagos értéke 47,49%, vagyis a helyesen felismert szegmensek aránya átlagosan 52,51%.

2. táblázat: Az átlagos EER értéke az akusztikai paraméterek függvényében

Származtatott jellemzők	SP	MFCC	MSL	RSE
Átlagos EER (%)	52,03	50,84	47,49	52,36

A második legjobban teljesítő jellemző az MFCC volt. Ennek átlagos EER-értéke 50,84% volt. Elmondható tehát az, hogy átlagosan 3,35%-os hibacsökkenést tudunk elérni a MSL-jellemző alkalmazásával az MFCC-vel elért eredményhez képest. Ez a javulás szignifikáns (Wilcoxon-próba: $Z = -2,211$; $p = 0,023$).

Megvizsgáltuk, hogy az EER értéke hogyan függ a jellemzők és a harmadik rétegben használt neuronok számától. Az eredmények azt mutatják, hogy a legjobb eredményt akkor kapjuk, ha MSL-jellemzőt és 500 neuront használunk a H3-ban (7. ábra).



7. ábra

Az EER értéke a jellemzők és a H3-ban lévő neuronok számának függvényében

A statisztikai elemzések alátámasztják, hogy a MSL szignifikánsan jobban teljesít attól függetlenül, hogy hány neuront használunk a harmadik rétegben (3. táblázat): MSL-MFCC: $Z = -2,201$; $p = 0,028$; MSL-SP: $Z = -2,201$; $p = 0,028$; MSL-RSE: $Z = -2,201$; $p = 0,028$.

3. táblázat: Az EER értéke a jellemzők és a H3-ban alkalmazott neuronok számának függvényében

	Neuronok száma a H3 rétegben	Akusztikai jellemzők			
		SP	MFCC	MSL	RSE
EER (%)	300	48,00	47,31	45,65	51,27
	400	52,45	50,12	44,87	52,15
	500	51,22	49,44	44,33	50,45
	600	50,05	48,02	47,48	51,81
	700	53,68	53,41	50,01	52,91
	800	56,77	56,76	52,59	55,58
	900	52,03	50,84	47,49	52,36

Az EER-értékekből azt látszik, hogy két esetben (SP és MFCC) akkor volt a legkisebb a hiba értéke, ha a harmadik rétegben 300 neuront használtunk. Az MSL és a RSE esetében pedig a legkisebb hibát akkor kaptuk, ha a neuronok száma 500 volt a harmadik rétegben. Általánosságban azonban az elmondható, hogy 500 neuron felett mindegyik jellemző esetében nőtt az EER értéke.

Az elért eredményeinket visszaellenőrizve elemeztük a hibák tulajdonságait. Az első és legnagyobb hibaforrás maga a kézi címkézés volt. Az egyszerre beszélések címkézése ugyanis sokszor igen nehéz feladat. A második hibaforrás a háttéracsatorna-jelzésre vezethető vissza; a legtöbb hibát (38,28%-ot) ezek okozták. Ez a nagyszámú hiba annak tudható be, hogy a háttéracsatorna-jelzések időtartama igen rövid, akár 60 ms-os is lehet, ami nem teszi lehetővé az elégséges számú jellemző kinyerését, így a belőlük származtatott statisztikai mutatók sem megbízhatók.

A háttéracsatorna-jelzések után a nevetés volt az a jelenség, amely rontotta az osztályozás eredményét. Az ilyen típusú hibák aránya 10,34% volt. Ennél a hibánál is jól látható, hogy a nevetés közben az akusztikumban igen erős torzulás jelenik meg, sokszor a felvétel túlvezéreltté válik, így az akusztikai jellemzőkinyerés nehezített.

Az egyszerrebeszélés-detektáló implementálása a beszélődetektálóba A beszélődetektáló felépítése

A beszélődetektáló rendszerben megvalósítottunk egy beszélőszegmentáló és egy beszélőklaszterező eljárást. A beszélőszegmentáló algoritmus kétlépcsős. Az első lépés egy BIC-alapú szegmentáló, amely MFCC-ket használ 2,5–3,5 kHz-es tartományban. A második lépésben pedig a téves riasztások kompenzálására egy Kullback–Leibler-divergencia alapú szegmentálót hoztunk létre, amelynek szintén az MFCC_{2,5–3,5} volt a bemeneti jellemzője. A beszédfordulók detektálása után az egyes beszédsegmentumokat beszélőkhöz rendeltünk, ami beszélőklaszterezési feladat. A klaszterezés bemeneti jellemzője a GMM-UBM-PCA (Gauss-keverék modell, általános háttérmodell és

főkomponens-analízis) *i*-vektorok. A klaszterezést pedig BIC-alapú nemellenőrzött módszerrel végeztünk el.

Eredmények

Az átfedő beszédrészek automatikus detektációjával átlagosan 2,5%-os relatív javulást tudtunk elérni, vagyis a DER értékét 31,21%-ról le tudtuk csökkenteni 28,71%-ra (4. táblázat). Ez a javulás szignifikáns (Wilcoxon-próba Monte Carlo szimulációval kiegészítve: $Z = -3,06$; $p = 0,002$).

4. táblázat: A DER értéke az átfedőbeszéd-detektálással és anélkül

Felvétel sorszáma	DER		Δ DER	Az átfedő beszéd és a társalgás hosszának aránya
	Átfedő beszéd tartalmaz	nem tartalmaz		
bea071n037	14,98%	12,36%	-2,62%	21,52%
bea072n038	27,83%	24,85%	-2,98%	38,62%
bea073n039	35,64%	33,70%	-1,94%	15,68%
bea074n040	23,89%	20,71%	-3,18%	44,28%
bea075n041	34,21%	32,79%	-1,42%	6,46%
bea094f039	33,63%	31,88%	-1,75%	13,39%
bea150n091	36,26%	34,60%	-1,66%	28,96%
bea166f066	27,74%	25,59%	-2,15%	31,67%
bea174n105	36,37%	33,31%	-3,06%	40,26%
bea184n111	35,07%	30,69%	-4,38%	38,99%
bea189n114	37,11%	33,55%	-3,56%	42,53%
bea192f077	31,80%	30,54%	-1,26%	40,66%
Átlagos	31,21%	28,71%	-2,5%	30,94%

Elemeztük, hogy a teszteléskor használt társalgásokban milyen arányban fordulnak elő egyszerre beszélések (4. táblázat). A táblázatban látható, hogy elég gyakoriak az átfedő részek ezen felvételekben. Jóllehet az egyszerre beszéléseket detektáló algoritmus eredményei nem voltak túl magasak, mégis statisztikailag igazolható relatív javulást tudtunk elérni a beszélődetektálással való implementációjával.

Következtetések

Az egyszerre beszélések magas, 12%-os előfordulása a korpuszban indokolja, hogy a beszélődetektálásban foglalkozzunk ezen jelenség automatikus osztályozásának lehetőségével. Jóllehet az egyszerre beszélések automatikus osztályozása igen fontos feladat a beszélődetektálásban, mégis csak néhány tanulmány foglalkozik ezzel a kérdéssel (pl. Mowlae et al. 2010; Saeidi et al. 2010). Boakye és munkatársai (2008) az AMI korpuszon (amely 18%-ban tartalmaz átfedő beszédet) 38%-os F-score-t értek el az átfedőbeszéd-detektálásra. Yella és Valente (2012) munkájukban azt a jelenséget igyekeztek mo-

dellezni, hogy a társalgásokban az átfedő beszédek előtt rövidebb a szünet (szüneteloszlás modellezése), mint a beszélőváltáskor. Az ezt modellező (HMM/GMM) módszerrel a beszélődetektálás DER értékét 8%-kal tudták csökkenteni. Prosódiai jellemzőket is tartalmazó eljárással Zelenák és Hernando (2011) hasonló F-score-t tudtak elérni az átfedőbeszéd-detektálásra, amely közel 40% volt. Vipperla és munkatársai (2012) konvolúciós nem-negatív ritka kódolással (convolutive non-negative sparse coding) az átfedőbeszéd-detektálásra 16,1%-os fedést és 28%-os pontosságot tudtak elérni a NIST RT korpuszon. Telefonbeszélgetésekre Ben-Harush és munkatársai (2009) az időtartományban adott entrópia jellemzők becslésével próbálta meg detektálni az egyszerre beszéléseket, ez a munka azonban csak kétbeszélős társalgásokra vonatkozik.

Yella és Bourlard (2013) Shriberg és munkatársainak 2001-es kutatási eredményeiből indulnak ki, amely azt a megfigyelést írta le, hogy az átfedő beszédresek előfordulása jóval gyakoribb a társalgások egy bizonyos részén. A megfigyelés arra is kiterjedt, hogy az átfedő beszéd megjelenése összefügg a beszédfordulók számával. Ezt a jelenséget kihasználva Yella és Bourlard létrehozta egy olyan algoritmust, amely ezt a jelenséget modellezi. Az általuk javasolt egyszerrebeszélés-detektálót beépítették beszélődetektálóba, amellyel 5%-os relatív DER-javulást tudtak elérni.

A fent leírt eredményekből látszik, hogy habár az egyszerre beszélések detektálásának eredménye jóval elmarad a kívánttól, a beszélődetektálóba való integrációja során a DER értéke csökkenthető.

Mivel sem az akusztikai jellemzőben, sem a detektáló algoritmus típusában nincs megegyezés, hogy melyik alkalmas az egyszerre beszélések detektálására, ezért a jelen kutatásban több akusztikai jellemzőt is teszteltünk, illetve egy olyan hibrid osztályozót hoztunk létre (DBN/SVM), amelyet igen hatékonyan alkalmaztak már más típusú problémák megoldására (Tang 2008).

A jelen kutatás során a legjobb eredményt a Mel-skála szerinti logaritmikus szűrőbank jellemző adta. Ez korrelál más kutatásokban is ezt a jellemzőt használó algoritmusok által elért eredménnyel, például a beszédhang-felismerésben (Li et al. 2012; Mohamed et al. 2012). Ezen tanulmányok arról számoltak be, hogy a Mel-skála szerinti logaritmikus szűrőbank jellemző jobban teljesített, mint az MFCC.

Teszteltük azt is, hogy hány neuront kell alkalmazni a harmadik rétegben. Az eredmények ebben a tekintetben azt mutatták, hogy 500 neuron után az EER értéke növekszik. A legjobb eredményt akkor kaptuk, ha Mel-skála szerinti logaritmikus szűrőbank jellemzőt és H1(300)-H2(600)-H3(500) topológijú DBN-t használtunk előfeldolgozásként, és SVM-RBF-et osztályozóként. Az EER értéke ekkor 44,33% volt. Kimutattuk, hogy a mély neuronháló alkalmasak a jelen problémában a jellemzők kialakítására nemellenőrzött tanulási folyamattal.

Eredményeink alapján kimutattuk, hogy ebben a feladatban nehézségeket okoznak a háttéracsatorna-jelzések és a nevetések, mivel ezek eredményezték a hibák többségét. Megjegyezzük viszont, hogy számos gyakorlati alkalmazás szempontjából – például ha az egyszerrebeszélés-detektálót beszédfelismerő előtt alkalmazzuk szűrőként a VAD kiegészítésére – kifejezetten előnyös lehet, ha az egyszerre beszélések mellett más, a felismerés kivitelezését lehetetlenné tévő események – így például a nevetés vagy bizonyos háttéracsatorna-jelzések – is detektálhatók (Neuberger–Beke 2013). Ebben az esetben az EER értéke jóval alacsonyabb lehet. Az egyszerre beszélés és egyéb események esetleges elkülönítése további osztályozással is megvalósítható, erre azonban jelen munkában nem térünk ki.

Az előzetes feltételezésüknek megfelelően, ha az egyszerre beszéléseket detektáló algoritmust integráltuk a beszélődetektálóba, akkor annak DER-értékét csökkenteni tudtuk, vagyis a beszélődetektáló eredményei javultak. Az átfedő beszéd automatikus detektációjával a DER értékét 31,21%-ról 28,71%-ra tudtuk csökkenteni, így átlagosan 2,5%-os relatív javulást lehetett elérni.

Összességében tehát elmondható, hogy bár az egyszerre beszélések detektálása az általunk kialakított módszerrel még mindig az elméletileg lehetséges értéknél alacsonyabb eredménnyel működik, mégis alkalmas arra, hogy a beszélődetektálóba integrálva növelje annak eredményességét.

Irodalom

- Bata Sarolta 2009. Beszélőváltások a beszédpartnerek személyes kapcsolatának függvényében. *Beszéd kutatás* 2009. 107–120.
- Beattie, Geoffrey 1983. *Talk: An analysis of speech and non-verbal behaviour in conversation*. Open University Press, Milton Keynes.
- Bellili, Abdel – Giloux, Michel – Gallinari, Patrick. 2001. An hybrid MLP-SVM handwritten digit recognizer. In: *Proceedings of the International Conference on Document Analysis and Recognition (ICDAR)*. 28–32.
- Ben-Harush, Oshry – Guterman, Hugo – Lapidot, Itzhak 2009. Frame level entropy based overlapped speech detection as a pre-processing stage for speaker diarization. In: *Machine Learning for Signal Processing, MLSP 2009. IEEE International Workshop*. 1–6.
- Boakye, Kofi A. 2008. *Audio segmentation for meetings speech processing*. PhD dissertation, University of California, Berkeley.
- Boakye, Kofi A. – Vinyals, Oriol – Friedland, Gerald 2011. Improved overlapped speech handling for speaker diarization. In: *Proceeding of Interspeech 2011*. Firenze, Italy. 941–944.
- Boakye, Kofi A. – Trueba-Hornero, Beatriz – Vinyals, Oriol – Friedland, Gerlad 2008. Overlapped speech detection for improved speaker diarization in multiparty meetings. In *Proceeding of ICASSP*. 4353–4356.
- Boronkai Dóra 2008a. Konverzációelemzés és anyanyelvtanítás I. *Anyanyelv-pedagógia* 2008/2. <http://www.anyanyelv-pedagogia.hu/cikkek.php?id=60>.

- Boronkai Dóra 2008b. Konverzációelemzés és anyanyelvtanítás II. *Anyanyelv-pedagógia* 2008/3–4. <http://www.anyanyelv-pedagogia.hu/cikkek.php?id=115>.
- Bourlard, Hervé – Morgan, Nelson 1993. Continuous speech recognition by connectionist statistical methods. *IEEE Transactions on Neural Networks* 4/6. 893–909.
- Burges, Christopher J. C. 1998. A Tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery* 2. 121–167.
- Cetin, Ozgür – Shriberg, Elizabeth 2006. Analysis of overlaps in meetings by dialog factors. Hot spots, speakers, and collection site: Insights for automatic speech recognition. In: *Proceedings of ICSLP*, Pittsburgh. 293–296.
- Chang, Chih-Chung – Lin, Chih-Jen 2012. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology* 2/3. 27. <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- Cho, KyungHyun 2014. *Advances in deep learning (draft)*. PhD dissertation. Aalto University, Aalto.
- Cho, KyungHyun. The RBM code for Matlab. <http://users.ics.tkk.fi/kcho/>.
- Dahl, George. E. – Ranzato, Marc’Aurelio R. – Mohamed, Abdel-rahman – Hinton, Geoffrey 2010. Phone recognition with the mean-covariance restricted Boltzmann machine. In: *NIPS 2010*. 469–477.
- De Rulter, Jan P. – Mitterer, Holger – Enfield, N. J. 2006. Predicting the end of a speakers turn: A cognitive cornerstone of conversation. *Language* 82/3. 515–535.
- Ellis, Dan 2005. Reproducing the feature outputs of common programs using Matlab and melfcc.m. <http://labrosa.ee.columbia.edu/matlab/rastamat/mfccs.html>
- Gósy, Mária 2012. BEA - A multifunctional Hungarian spoken language database. *The Phonician* 105–106. 50–61.
- Grácsi Tekla Etelka – Bata Sarolta 2010. Megszólalási formák és funkciók az összeszokottság függvényében. In Geccső Tamás – Sárdi Csilla (szerk.): *Új módszerek az alkalmazott nyelvészeti kutatásban*. Kodolányi János Főiskola–Tinta Könyvkiadó, Székesfehérvár, Budapest. 28–32.
- Hámori Ágnes 2006. A társalgási műfajokról. In Tolcsvai Nagy Gábor (szerk.): *Szöveg és típus. Szövegtipológiai tanulmányok*. Tinta Kiadó, Budapest. 157–181.
- Heldner, Mattias – Edlund, Jens 2010. Pauses, gaps and overlaps in conversations. *Journal of Phonetics* 38. 555–568.
- Hsu, Chih-Wei – Chang, Chih-Chung – Lin, Chih-Jen 2003. *A practical guide to support vector classification*. Technical report, Department of Computer Science, National Taiwan University, Taipei.
- Iványi Zsuzsanna 2001. A nyelvészeti konverzációelemzés. *Magyar Nyelvőr* 125. 74–93.
- Li, Jinyu – Yu, Dong – Huang, Jui-Ting – Gong, Yifan 2012. Improving wideband speech recognition using mixed-bandwidth training data in CD-DNN-HMM. In: *Proceeding IEEE Workshop on Spoken Language Technology*. 131–136.
- Markó Alexandra 2005. *A spontán beszéd néhány szuprasegmentális jellegzetessége. Monologikus és dialogikus szövegek összevetése, valamint a hűmmögés vizsgálata*. PhD-értekezés. ELTE, Budapest.
- Martin, Alvin F. – Doddington, George R. – Kamm, Terri – Ordowski, Mark – Przybocki, Mark A. 1997. The DET curve in assessment of detection task performance. In: *Proceedings of Eurospeech*. Rhodes, Greece. 1899–1903.

- Moattar, Hossein – Homayounpour, Mohammed M. 2006. Speech overlap detection using spectral features and its application in speech indexing. *Information and Communication Technologies* 2/1. 1270–1274.
- Mohamed, Abdel-rahman – Hinton, Geoffrey – Penn, Gerald 2012. Understanding how deep belief networks perform acoustic modelling. In: *Proceedings of ICASSP*. 4273–4276.
- Mowlae, Pejman – Christensen, Mads G. – Tan, Zheng-Hua – Jensen, Søren H. 2010. A MAP criterion for detecting the number of speakers at frame level in model-based single-channel speech separation. In: *Signals, Systems and Computers (ASILOMAR)*. 538–541.
- Neuberger, Tilda – Beke, András 2013. Automatic laughter detection in spontaneous speech using GMM-SVM method. In: *Proceedings of TSD2013*, Springer Berlin Heidelberg. 113–120.
- Ogden, Richard 2004. Non-modal voice quality and turn-taking in Finnish. In Couper-Kuhlen, Elisabeth – Ford, Cecilia E. (eds.): *Sound patterns in interaction*. Benjamins, Amsterdam. 29–62.
- Otterson, Scott – Ostendorf, Mari 2007. Efficient use of overlap information in speaker diarization. In: *Proceedings of ASRU*. Kyoto, Japan. 683–686.
- Sacks, Harvey – Schegloff, Emanuel A. – Jefferson, Gail 1974. A simplest systematics for the organization of turntaking for conversation. *Language* 50. 696–735.
- Saeidi, Rahim – Mowlae, Pejman – Kinnunen, Tom – Tan, Zheng-Hua – Christensen, Mads G. – Jensen, Søren H. – Franti, Pasi 2010. Improving monaural speaker identification by double-talk detection. In: *Eleventh Annual Conference of the International Speech Communication Association*. 1069–1072.
- Sarikaya, Ruhi – Pellom, Bryan L. – Hansen, John H. L. 1998. Wavelet packet transform features with application to speaker identification. In: *Proceedings of IEEE Nordic Signal processing Symposium*. Visgo, Denmark. 81–84.
- Shriberg, Elizabeth – Stolcke, Andreas – Baron, Don 2001. Observations on Overlap: Findings and Implications for Automatic Processing of Multi-Party Conversation. In: *Proceedings of EUROSPEECH*. Aalborg, Denmark. 1359–1362.
- Suykens, Johan A. K. – Van Gestel, Tony – De Brabanter, Jos – De Moor, Bart – Vandewalle, Joos 2002. Least squares support vector machines. *World Scientific* 4. Singapore.
- Tang, Huixuan 2008. A comparative evaluation of deep belief nets in semi-supervised learning. In *Report for CSC2515*. http://www.cs.toronto.edu/~hxtang/projects/dbn_eval/dbn_eval.pdf
- Tóth, László – Grósz, Tamás 2013. A comparison of deep neural network training methods for large vocabulary speech recognition. In: *Proceedings of TSD2013*. Springer Berlin Heidelberg. 36–43.
- Trueba-Hornero, Beatriz 2008 *Handling overlapped speech in speaker diarization*. Master's thesis. Universitat Politècnica de Catalunya, Barcelona.
- Vipperla, Ravichander – Geiger, Jürgen T. – Bozonnet, Simon – Wang, Dong – Evans, Nicholas – Schuller, Björn – Rigoll, Gerhard 2012. Speech overlap detection and attribution using convolutive non-negative sparse coding. In: *Proceedings of ICASSP-12*. 4181–4184.
- Wells, Bill – Peppé, Sue 1996. Ending up in Ulster: Prosody and turn-taking in English dialects. In Couper-Kuhlen, Elisabeth – Selting, Margret (eds.): *Prosody in*

- Conversation: Interactional studies*. Cambridge University Press, Cambridge–New York–Melbourne. 101–130.
- Wooters, Chuck – Huijberts, Marijn 2007. The ICSI RT07s speaker diarization system. In *Proceedings of the Rich Transcription, Meeting Recognition Evaluation Workshop*. Baltimore, MD.
- Yella, Sree Harsha – Bourlard, Hervé 2013. Improved overlap speech diarization of meeting recordings using long-term conversational features. In: *Proceedings of ICASSP*. 7746–7750.
- Yella, Sree Harsha – Valente, Fabio 2012. Speaker diarization of overlapping speech based on silence distribution in meetings recordings. In: *Proceedings of Interspeech 2012*. Portland, USA.
- Zelenák, Martin – Hernando, Javier 2011. The detection of overlapping speech with prosodic features for speaker diarization. In: *Proceedings of Interspeech 2011*. 32–35.
- Zelenák, Martin – Segura, Carlos – Hernando, Javier 2010. Overlap detection for speaker diarization by fusing spectral and spatial features. In: *Proceedings of Interspeech 2010*, Makuhari, Japan. 2302–2305.

A kutatás a 108762-es számú OTKA-pályázat keretében készült.

IRREGULÁRIS BESZÉD REGULÁRISSÁ ALAKÍTÁSA BESZÉDKÓDOLÁSON ALAPULÓ MÓDSZERREL

Csapó Tamás Gábor – Németh Géza

Bevezetés

Az emberi beszédben a zöngés hangok képzésekor a hangszalagok általában közel tökéletesen, kváziperiodikusan rezegnek, azaz az egyes zöngeperiódusok között csak kis változások figyelhetők meg. Ezt reguláris, modális, ideális zöngképzésnek nevezzük. A természetes beszéd során azonban hosszabb-rövidebb időtartamra irregulárisra válhat a hangszalagok rezgése. Ez eltér a modális zöngképzéstől, és ezért ezt irreguláris fonációnak, glottalizációnak, érdes zöngének vagy recsegő beszédnek nevezik (Gósy 2004: 30; Markó 2005: 61; Böhm 2009). A jelenség a zöngeperiódusok hosszának és/vagy amplitúdójának hirtelen megváltozásából adódik. Az irreguláris fonáció előfordul egészséges és patológus beszélők esetén is (Redi-Shattuck-Hufnagel 2001), általában szakaszhatárokon (pl. mondat végén) (Markó 2011) vagy magánhangzó-magánhangzó kapcsolatokban (Markó 2012a; Markó 2012b). Gyakran kíséri extrém alacsony alaphangfrekvencia és a glottális pulzusok hirtelen változása (Blomgren et al. 1998). Érzetileg recsegő, rekedtes, érdes jellegű beszédet jelent (Fónagy–Magdics 1967; Gósy 2004). Ha az irreguláris fonáció rövid ideig van jelen a beszédben (néhány periódus), akkor nem is nagyon érzékeljük, tehát azt lehet mondani, hogy az ilyen irreguláris zöngképzés a legtöbb esetben nem zavarja a hallgatót.

A glottalizáció előfordulása függ a prozódiai szerkezettől [gyakran egybeesik prozódiai határokkal, például szünetekkel (Slifka 2007) és hangsúlyos szótagokkal (Dilley et al. 1996)], valamint információt hordoz a beszélő személyéről, nyelvjárásáról, hangulatáról, érzelmi állapotáról, és arról, hogy a hangszalagok egészségesek-e (Gobl–Chasaide 2003; Malyska 2008). A glottalizáció akár a beszédhangok 15%-ában is előfordulhat egy-egy beszélő esetén, így egyáltalán nem elhanyagolható jelenség (Böhm et al. 2009). A legtöbb beszédtechnológiai módszert idealizált beszéd feldolgozására készítették. Az irreguláris fonáció problémákat okozhat a beszédanalízis-módszerekben (pl. f_0 -mérés és spektrális analízis). A fentiek miatt az irreguláris zöngével képzett beszéd megfelelő modellezése hozzájárulhat a pontosabb beszédanalízis-módszerekhez és természetesebb hangzású beszéd szintetizátor rendszerek elkészítéséhez.

Ugyan már léteznek módszerek a glottalizáció jelenségének elemzésére (Markó 2005; Böhm 2009), detektálására és osztályozására (Böhm et al. 2009; Beke–Heltovics 2010; Kane et al. 2013), reguláris-irreguláris transz-

formációjára (Böhm et al. 2008; Ruinskiy–Lavner 2008), valamint az irreguláris zöngével kiegészített beszéd szintézisre (Silén et al. 2009; Raitio et al. 2013; Csapó–Németh 2013), de az irreguláris zöngéképzés gépi javításával, vagyis az irreguláris-reguláris transzformációval eddig keveset foglalkoztak.

A szakirodalomban számos beszéd kódoló módszerről olvashatunk, melyeknek eredeti célja a beszéd paramétereinek bontása (kódolás, analízis lépés) azért, hogy a távközlési csatornán minél kisebb sáv szélesség mellett lehessen átvinni a jelet (beszédet) (Németh–Olaszy 2010: 244). Az átvitel után, a vevő oldalon a paramétereket visszaalakítják beszédjellé (dekódolás, szintézis lépés). A kódolási technikákat három csoportba lehet osztani: hullámforma-kódolás, parametrikus vagy forráskódolás, valamint hibrid kódolás. A hullámforma-kódolás tetszőleges sávkorlátozott jel digitális tárolására alkalmas, és a jel redundanciájának csökkentésével törekszik az alakhűség megtartására. A parametrikus kódolás esetén forrásmodell (beszédkeltési modell) alkalmaznak, ami miatt ez a kódolási forma csak beszédjelle alkalmazható. A parametrikus kódolók családjába tartozik az LPC (Linear Predictive Coding) kódoló, valamint ennek továbbfejlesztett változatai, melyek az elsődleges cél mellett alkalmasak a beszédjel tulajdonságainak változtatására is (pl. f_0 -módosítás). A hibrid kódolás az előző kettő előnyeit ötvözi. A jelen kutatáshoz egy újonnan kifejlesztett parametrikus kódolót használunk fel (Csapó–Németh 2012).

A tanulmányban bemutatunk egy új eljárást, amely irreguláris beszéd regulárisra alakítására alkalmas egy jelfeldolgozási módszerrel. A beszéd kódoló analízis-szintézis felbontáson alapul. A módszer a forrás-szűrő modell felhasználásával az eredeti glottalizált beszédet szétbontja gerjesztési és spektrális komponensekre (analízis), melyek a glottális forrást, illetve az artikulációs csatornát modellezik. Az analízis során kapott paraméterek (pl. f_0) értékeit a glottalizált szakaszokon javítjuk, majd a visszaalakításkor előállítjuk a javított beszédet (szintézis). A tanulmány további részében percepció és akusztikai tesztekben vizsgáljuk a módszer eredményességét.

Anyag, módszer, kísérleti személyek

Az irreguláris zöngéképzéssel kapcsolatos kísérleteket a PPBA adatbázisból kiválasztott négy magyar anyanyelvű beszélő felolvasott anyagán végeztük (Olaszy 2013). Három férfitől (FF1, FF3 és FF4) és egy nőtől (NO3) származó, professzionális körülmények között rögzített, 44,1 kHz-es mintavételezéssel 16 biten digitalizált hangfelvételt használtunk fel. Az adatbázis beszélőnként közel ugyanazt az 1940 mondatot tartalmazza, amely nagyjából 2 órányi tiszta hangfelvételt jelent. Az adatbázishoz szöveges címkézés, automatikus fonetikus átírat és ennek kézzel javított változata, valamint hang- és szóhatárjelölés is tartozik. A precíziós annotálás eredményeként az adatbázis megbízható: a hanghatárok pontossága 10 ms-on belül van, és a fonetikai átírat pontosan megfelel a hanghullámnak.

Az irreguláris beszéd regulárisá alakításának módszere

A jelen tanulmányban bemutatott módszer a korábbi Csapó–Németh (2012) analízis-szintézis módszert (mely egy forrás-szűrő szétválasztáson alapuló gerjesztési modell) egészíti ki egy olyan jelfeldolgozó eljárással, amely alkalmas a glottalizált beszéd modálissá alakítására, tehát az irreguláris zöngéképzés gépi javítására. Az új eljárás három fő lépésből áll: analízis, transzformáció és szintézis. Az analízis hasonlóan történik, mint a Csapó–Németh (2012) gerjesztési modellben, azaz az eredeti beszédjelből spektrális paramétereket, majd a lineáris predikció alapú inverz szűréssel származtatott maradékjelből gerjesztési paramétereket mérünk. A gerjesztési paraméterek modellezik a glottális forrást, a spektrális paraméterek pedig az artikulációs csatornát. Az analízis után a transzformáció lépésben a paramétereket módosítjuk, majd a gerjesztési modell szintézisével a paraméterekből visszaállítjuk a javított beszédjelet.

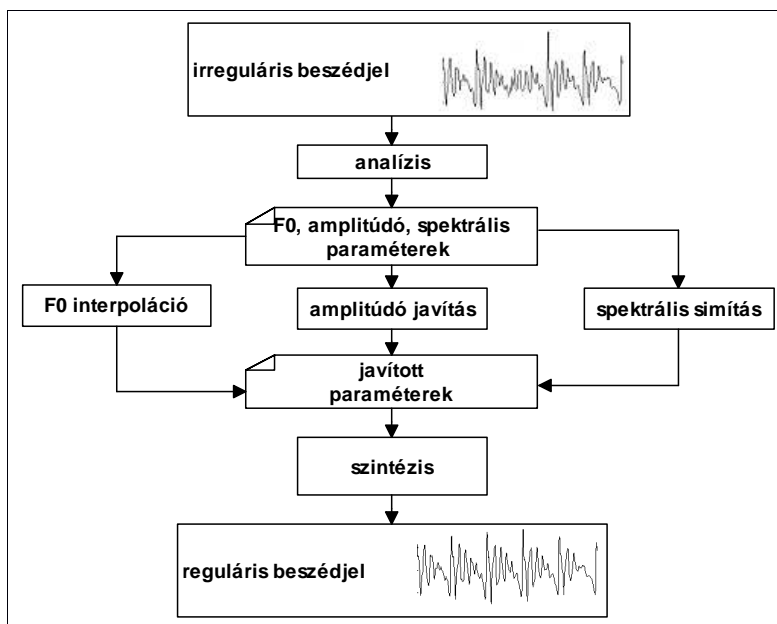
A transzformáció során az eredeti beszédből származtatott maradékjelnek csak azon szakaszait vizsgáljuk, amelyet irreguláriszöngé-címkék jeleznek, a modális zöngés és zöngétlen maradékjel részeket változtatlanul hagyjuk. A transzformációs eljárás működését az 1. ábra ismerteti.

A gerjesztési modell analízis lépésének eredményeként kapott paraméterek, melyeket 50 ms-os keretekre számolunk:

- **f_0** : a keret alaphérvenciája a Snack RAPT algoritmussal számítva (Talkin 1995),
- **gain**: a keret energiája átlagos négyzetes középként számítva,
- **MGC**: a keret spektrális paraméterei Mel-általánosított kepsztrum (Mel-Generalized Cepstrum) alapon (Tokuda et al. 1994).

Az analízis eredményeként kapott f_0 -értékeket interpoláljuk, míg a gain- és MGC-értékeket simítjuk az irreguláris szakaszokon. Az analízis során a glottalizáció megzavarhatja az f_0 -detekciót: a hirtelen alaphérvencia- és amplitúdóváltozás miatt előfordulhat, hogy egy egyébként nyelvileg zöngésnek várt keretet (például magánhangzóban) zöngétlennek jelöl a detektor, vagy az eredeti érték felét méri. Ezeket a helyeket megjelöljük, majd a mért f_0 -menetet interpoláljuk (korrigáljuk az f_0 -értékeket) azokban a zöngés szakaszokban, ahol az algoritmus nem detektált zöngét. Ezen szakaszokat a beszédmintához tartozó hanghatárjelölés alapján határozzuk meg. Ha egy magánhangzóban egyáltalán nem mért f_0 -értéket a zöngedetektor, akkor a mondat dallamának megfelelő ereszkedő f_0 -menetet állítunk be. A kísérletek során minden f_0 -menetet kézzel ellenőriztünk és javítottunk, emiatt a módszer félautomatikus működésű. Az irreguláris fonáció kis perturbációkat okoz a keretenkénti gain- és MGC-értékekben az irreguláris zöngeperiódusok amplitúdójának hirtelen változása miatt. Emiatt 5 pontos simítást végeztünk ezeken a paramétereken, amely tapasztalataink szerint megfelelőnek bizonyult a perturbációk eltüntetésére. A szintézis további lépései megegyeznek a Csapó–Németh (2012) gerjesztési modellben ismertetett lépésekkel, azaz a paramétereknek

megfelelő maradékjel-elemeket keresünk a kódkönyvből a célköltség és összefüzési költség felhasználásával, majd ezeket átlapolt összeadással összefűzzük. A zöngés és zöngétlen részeket egyesítve az energia megfelelő beállítása után spektrális szűréssel kapjuk meg a transzformáció kimeneti beszédjelét.



1. ábra

Az irreguláris-reguláris transzformáció működése. Az eljárás a Csapó–Németh (2012) maradékjel-kódkönyv alapú gerjesztési modellt alkalmazza

Percepciós teszt

A transzformációs eljárás eredményességét internetes percepció (meghallgatásos) kísérlettel vizsgáltuk. A kísérletben a tesztelők az egyes hangminták meghallgatása után 1–5 skálás MOS (Mean Opinion Score) jellegű kérdésekre válaszoltak. A teszt elején a kísérleti személyek egy ismertetőt olvashattak a kísérlet témájáról és menetéről, majd néhány adat (nem, kor, esz-köz, beszédtechnológiai ismeretek) megadását kértük tőlük.

Az irreguláris-reguláris transzformáció működését a PPBA adatbázis négy beszélőjének (3 férfi: FF1, FF3 és FF4 és egy nő: NO3) hanganyagán teszteltük (Olaszy 2013). Mind a négy beszélő gyakran használ irreguláris fonációt, elsősorban szakaszhatárokon (pl. a mondatok végén). Kiválasztottunk 4-4 szót, amelyek reguláris és irreguláris formában is előfordultak az adatbázisban. Ezután az irreguláris változatot transzformáltuk a fenti módszerrel. Bi-

zonyos esetekben csak egy magánhangzó volt glottalizált, míg más mintákban a teljes zöngés szakaszt irreguláris módon ejtette a beszélő. Azokban az esetekben, amelyekben az f_0 -interpoláció nem volt megfelelő, kézzel javítottuk az f_0 -menetet, hogy illeszkedjen a mondat ereszkedő dallammenetéhez. A szavak 3-3 változatát (eredeti irreguláris, regulárisá transzformált és eredeti reguláris) meghallgatásos tesztben hasonlítottuk össze. A 3. ábra egy példát mutat a teszt egyik szavának 3-3 változatára.

Az internetes meghallgatásos tesztben összesen 48 szót kellett értékelni (4 beszélő \times 4 szó \times 3 változat) természetesség és érdekesség szerint. A teszt megkezdése előtt a kísérleti személyeket megkértük, hogy hallgassanak meg néhány glottalizált beszédmintát, hogy egyértelműsítsük az „érdes” kifejezés jelentését. A tesztelők minden minta meghallgatása után két MOS-jellegű kérdésre válaszoltak: 1) „Kérlek, jelöld be, hogy mennyire érzed érdesnek (glottalizáltnak) a hallott hangot!” („1 – egyáltalán nem érdes” ... „5 – nagyon érdes”); 2) „Kérlek, jelöld be, hogy mennyire érzed természetesnek a hallott hangot!” („1 – nagyon természetellenes” ... „5 – teljesen természetes”). A szavakat a tesztelők véletlen sorrendben hallgatták meg.

A tesztet összesen 9 magyar anyanyelvű tesztelő végezte el (mind a 9 férfi volt). Közülük hárman beszédkutatók voltak, míg a többiek egyetemi hallgatók. Az átlagos életkor 23,67 év volt (szórás: 3,20 év). Heten fejhallgatóval végezték a kísérletet, ketten hangszórón hallgatták a mintákat. Átlagosan 6,92 percig tartott a teszt kitöltése (szórás: 1,39 perc).

Akusztikai teszt

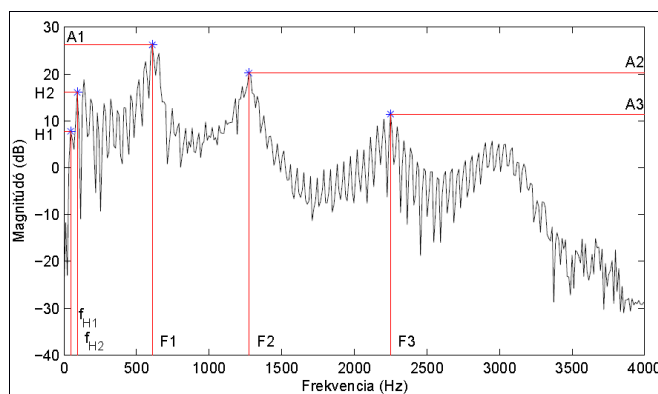
A meghallgatásos teszthez kiválasztott beszédmintákon akusztikai elemzést is végeztünk. A zöngeminőségnek számos akusztikai megfelelője van, melyeket a szakirodalomban következetesen használnak (Blomgren et al. 1998). Ez alapján megvizsgáltuk az eredeti irreguláris, regulárisá transzformált és eredeti reguláris beszédmintákat néhány akusztikai jegy szempontjából.

A szakirodalomból kiválasztottunk három olyan akusztikai jegyet, amelyeket korábban irreguláris és reguláris beszéd megkülönböztetésére használtak (Klatt–Klatt 1990; Böhm et al. 2008). Ezek alapján irreguláris zöngéképzés esetén a hangrés nyitott idejének aránya, vagyis a nyitott hányad (open quotient, OQ) alacsonyabb, mint reguláris zöngében. Az első formáns sávszélessége (first formant bandwidth, B1) nagyobb a glottalizáció során a gégeben előforduló nagyobb mértékű akusztikai veszteség miatt. Irreguláris zöngéképzés során a hangszalagok záródása összefüggéstelenebb, azaz a spektrális lejtés (spectral tilt, TL) meredekebb, mint reguláris beszédben. A transzformáció hatását az OQ, B1, TL akusztikai jellemzőkre mérésekkel vizsgáltuk. A méréseket spektrális tartományban végeztük, mivel így könnyebb a paraméterek számítása (Böhm 2009: 53). Holmberg és munkatársai (1995) kimutatták, hogy az OQ arányos az első és második harmonikus dB-ben mért különbségével (H_1-H_2), B1 fordítottan arányos H_1 és az első formáns ampli-

túdójának különbségével ($H_1 - A_1$), míg a TL korrelál H_1 és a harmadik formáns amplitúdójának különbségével ($H_1 - A_3$).

A H_1 , H_2 és A_3 értékeket a formánsok hatása befolyásolhatja, ezért az Iseli–Alwan (2004) által javasolt kompenzációt alkalmaztuk a VoiceSauce program implementációjában. Ez alapján a H_1 és H_2 értékét az első és második formáns frekvenciája szerint korrigáltuk (H_1^* és H_2^*), valamint az A_3 értékét az első három formáns frekvenciája szerint kompenzáltuk (A_3^*). Végül kiszámítottuk az amplitúdókülönbségeket: $H_1^* - H_2^*$, $H_1^* - A_1$ és $H_1^* - A_3^*$.

A paramétereket a meghallgatásos teszt mintáin mértük (eredeti irreguláris, reguláris transzformált és eredeti reguláris változatok). A hangfájlokat először 3,8 kHz-es aluláteresztő szűréssel engedjük át, majd újramintavételeztük 8 kHz-en, ami biztosította, hogy a spektrumban csak a 3,8 kHz alatti tartomány látható. Ezután az eredeti irreguláris változatokból kiválasztottunk egy glottalizált magánhangzót, majd ennek 3-3 pontját jelöltük ki (nagyjából egyenletesen elosztva és a zöngéhatároknak megfelelően). A Wavesurfer programot használva 512 pontos FFT-analízist végeztünk a Hanning-ablakozott jelen, majd vizuálisan leolvastuk a H_1 , H_2 amplitúdókat és ezek frekvenciáit, az F_1 , F_2 , F_3 , valamint A_1 és A_3 értékeket. Az irreguláris változatokban gyakran erős alharmonikusok jelentek meg; itt a H_1 és H_2 értékét a spektrális csúcsok közül az első kettőben mértük. A formánsok értékeit a legnagyobb spektrális csúcs frekvenciájaként és amplitúdójaként mértük. A mérés módszere a 2. ábrán látható.



2. ábra

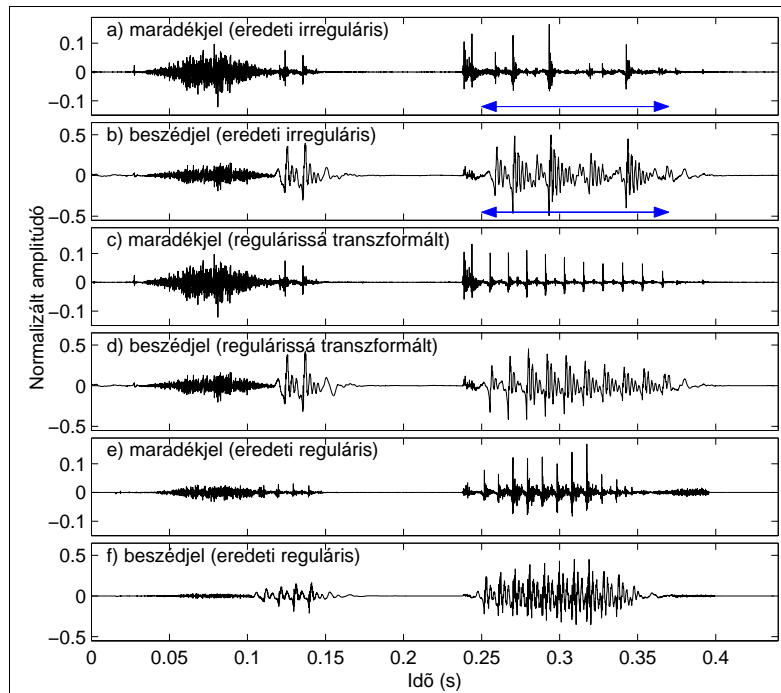
Az első két harmonikus (H_1 , H_2) és az első három formáns (F_1 , F_2 és F_3) frekvenciájának és amplitúdójának (A_1 , A_2 és A_3) mérése az FFT-spektrum alapján

A kutatás során a transzformációs eljárásokat a Matlab szoftverrel készítettük; a statisztikai teszteteket az SPSS programban végeztük. A statisztikai

elemzések során párosított mintás t -próbát, nem paraméteres összetartozó mintás Wilcoxon-tesztet és Tukey-HSD post hoc tesztel kiegészített egytényezős ANOVA elemzést alkalmaztunk a Matlab és SPSS programokkal. Az elemzések során kétoldalas $p < 0,05$ szignifikanciaszint alatt (95% konfidenciaszint felett) vetjük el a nullhipotézist.

Eredmények

A 3. ábra az irreguláris-reguláris transzformáció eredményére mutat egy példát. Az ábrán észrevehető, hogy a „regulárisra transzformált” (c és d) és az „eredeti reguláris” (e és f) változatoknak hasonló zöngeperiódusai vannak, míg az „eredeti irreguláris” (a és b) jel ettől lényegesen eltérő, és periódusonkénti amplitúdóingadozást tartalmaz. A c–e, illetve d–f ábrák közti különbségek azért fordulnak elő, mert ez a szó két különböző realizációja, így kis eltérések láthatóak az egyes beszédhangok időtartamában és átlagos amplitúdójában.

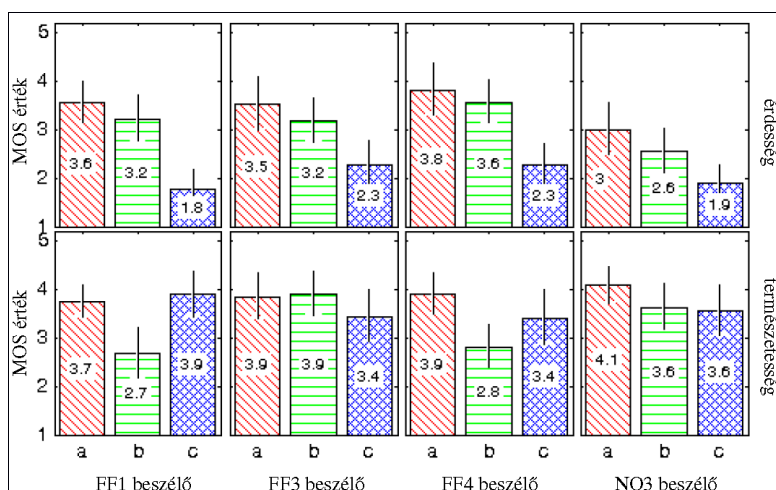


3. ábra

A kiejtett és transzformált „cipő” szó beszédhullámformái és maradékjelei FF3 beszélőtől. (Nyíl jelöli az irreguláris zöngét.)

Percepció teszt

A tesztelők értékelését a 4. ábra mutatja, melyet nem paraméteres összetartozó mintás Wilcoxon-próbával hasonlítottunk össze. Az „érdeesség” kérdésnél az „1” érték jelöli az egyáltalán nem érdekes, míg az „5” a nagyon érdekes válaszokat. A „természetesség” kérdésnél az „1” érték jelöli a nagyon természetellenes, míg az „5” a teljesen természetes válaszokat. Az elemzések szerint az eredeti irreguláris és a regulárisra transzformált minták nem különböznek szignifikánsan érdeesség szempontjából ($p = 0,13$), amennyiben az összes beszélőt egybe vesszük. Az ábra MOS-értékeit megvizsgálva azt vehetjük észre, hogy az eredeti irreguláris minták érdeessége csökkent, de a statisztikai próba ezt nem mutatta ki. Összességében a módszer szignifikáns mértékben csökkentette az eredeti minták természetességét ($p < 0,05$). A természetesség megőrzésében nagy beszélőnkénti különbség látható: az FF3 és NO3 beszélők esetében nem csökkent szignifikánsan a minták természetessége, míg az FF1 és FF4 beszélő esetén igen.



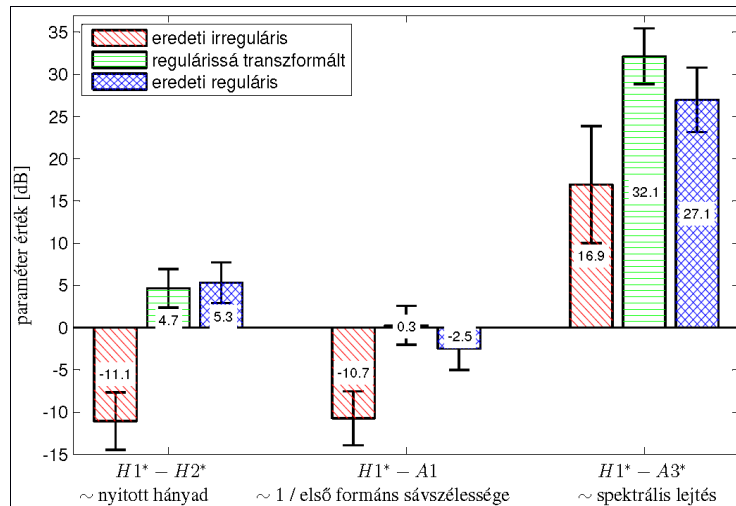
4. ábra

Az irreguláris-reguláris transzformációval módosított szavak percepció elemzésének eredménye: a) eredeti irreguláris, b) regulárisra transzformált, c) eredeti reguláris. A függőleges fekete vonalak a 95%-os konfidenciaintervallumot jelölik

A meghallgatásos teszt eredménye az, hogy az irreguláris-reguláris transzformációs módszer kismértékben csökkentette a beszédminták érzeti érdeességét, és a négyből két beszélő esetén ezt a természetesség csökkentése nélkül tudta megtenni.

Akusztikai teszt

A három mért akusztikai paramétert ($H_1^*-H_2^*$: nyitott hányad, $H_1^*-A_1$: első formáns sáv szélessége, $H_1^*-A_3^*$: spektrális lejtés) a három beszédminta-típuson az 5. ábra mutatja be. ANOVA-elemzést végeztünk, mely szerint a beszédminták típusának szignifikáns hatása volt mindhárom paraméterre [$F(2, 141) = 44,16; 17,26; 9,36$; sorban; $p < 0,001$]. Tukey-HSD post hoc teszttel hasonlítottuk össze a beszédmintatípusok átlagos paramétereinek értékét. Ez alapján megállapítottuk, hogy a $H_1^*-H_2^*$ megközelítőleg azonos az eredeti reguláris és a transzformált beszédrészleteken ($p = 0,938$, n.sz. különbség), míg szignifikánsan különböző az eredeti irreguláris mintákhoz képest ($p < 0,001$). A nyitott hányad szempontjából a transzformált változatok tehát közel vannak a modális beszédhez. Az irreguláris zöngével képzett szavak $H_1^*-A_1$ és $H_1^*-A_3^*$ különbségei szintén szignifikánsan különbözőek az eredeti reguláris és a transzformált változatokhoz képest ($p < 0,001$ és $p < 0,05$), de az eredeti reguláris és a transzformált változatokban közel meg-egyeznek ($p = 0,336$ és $p = 0,321$, n.sz. különbség). Eszerint a transzformált minták közel vannak az eredeti modális felvételekhez B_1 és TL tekintetében is. A spektrális lejtés szempontjából viszont a transzformált minták értékei jóval magasabbak a természetes mintákénál, ami magyarázhatja, hogy miért érezhették a meghallgatásos teszt résztvevői a minták egy részét kevésbé természetesnek.



5. ábra

Az irreguláris–reguláris transzformációval módosított szavak akusztikai elemzésének eredménye. A függőleges fekete vonalak a 95%-os konfidenciaintervallumot jelölik

A transzformációs eljárás a vizsgált akusztikai jegyek (nyitott hányad, első formáns sáv szélessége és spektrális lejtés) szempontjából tehát a reguláris beszédre jellemző értékek irányába módosítja az irreguláris beszédjelet.

Összefoglalás

A tanulmányban ismertettünk egy transzformációs eljárást, amely irreguláris-reguláris beszéd átalakítására alkalmas egy beszédkódoláson alapuló jel-feldolgozási módszerrel. A transzformáció eredményét percepciók kísérletben és akusztikai elemzés során vizsgáltuk. Ezek alapján az irregulárisból regulárisra alakított beszéd kevésbé érdekes, mint az eredeti minták; valamint három akusztikai jegy szempontjából közel van az eredeti reguláris beszédhez.

A beszédmintákat utólagosan megvizsgálva a két beszélőnél jelentkező természetesség csökkenését valószínűleg az alkalmazott Csapó–Németh (2012) gerjesztési modell okozta: a maradékjel-kódkönyvek felépítése és a transzformáció során az összefűzési költség számításának változtatásával várhatóan jobb eredmény érhető el. A szintézis lépésben az elemkiválasztásnál előfordult, hogy az összefűzési költség miatt ugyanazon maradékjel-periódus ismételtén megjelent egymás után többször, így robotos, gépies hangzást eredményezett. A személyfüggés oka az lehet, hogy a glottalizáció különböző megjelenési formái közül a transzformációs algoritmus valószínűleg nem minden esetben tudja megtartani az eredeti beszéd természetességét.

A percepciók teszt eredményeit az is befolyásolhatta, hogy bizonyos tesztlők ismerték a beszélők hangjait, míg mások nem. Bóhm–Ujváry (2008) kimutatta, hogy az irreguláris fonáció gyakorisága jellemző az egyes beszélőkre. A jövőben tervezzük új percepciók teszt végrehajtását hosszabb beszédmintákkal, nagyobb tesztlői létszámmal, valamint különválasztva a beszélőket ismerő és nem ismerő tesztlőket. Az akusztikai elemzés megbízhatóságát növelheti az eredeti reguláris és regulárisra transzformált minták között más egyezőségi mutató (pl. korrelációs együttható, átlagos négyzetes hiba) használata.

A kutatásban bemutatott glottalizációjavító eljárást ki lehet terjeszteni hosszabb beszédszakaszokra is, amivel rekedtes, patológus hangokat várhatóan szebbé, kellemesebbé lehet tenni beszédkozmetikai eljárással (pl. színészek, bemondók hangja). Az irreguláris-reguláris átalakító eljárás automatikussá kiegészített változatával beszédadatbázisokból el lehetne tüntetni az irreguláris zöngéjű szakaszokat, ezáltal ideálisabbá téve a beszédet a további feldolgozás és felhasználás céljából.

Irodalom

- Beke András – Heltovics Éva 2010. A glottalizált magánhangzók automatikus osztályozása spontán magyar beszédben. *Beszédkutatás 2010*. 253–263.
- Blomgren, Michael – Chen, Yang – Ng, Manwa L. – Gilbert, Harvey R. 1998. Acoustic, aerodynamic, physiologic, and perceptual properties of modal and vocal fry registers. *The Journal of the Acoustical Society of America* 103/5. 2649–2658.

- Bóhm, Tamás 2009. *Analysis and modeling of speech produced with irregular phonation*. PhD-disszertáció. BME TMIT, Budapest.
- Bóhm, Tamás – Audibert, Nicolas – Shattuck-Hufnagel, Stefanie – Németh, Géza – Aubergé, Véronique 2008. Transforming modal voice into irregular voice by amplitude scaling of individual glottal cycles. In: *Acoustics'08*. Paris, France. 6141–6146.
- Bóhm, Tamás – Both, Zoltán – Németh, Géza 2009. Automatic classification of regular vs. irregular phonation types. In: *NOLISP*. Vic, Spain. 43–50.
- Bóhm Tamás – Ujváry István 2008. Az irreguláris fonáció mint egyéni hangjellemző a magyar beszédben. *Beszédkutató 2008*. 108–120.
- Csapó, Tamás Gábor – Németh, Géza 2012. A novel codebook-based excitation model for use in speech synthesis. In: *CogInfoCom 2012*. Kosice, Slovakia. 661–665.
- Csapó, Tamás Gábor – Németh, Géza 2013. A novel irregular voice model for HMM-based speech synthesis. In: *Proceedings of ISCA SSW8*. Barcelona, Spain. 229–234.
- Dilley, Laura – Shattuck-Hufnagel, Stefanie – Ostendorf, Mari 1996. Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics* 24/4. 423–444.
- Fónagy Iván – Magdics Klára 1967. *A magyar beszéd dallama*. Akadémiai Kiadó, Budapest.
- Gobl, Christer – Ní Chasaide, Ailbhe 2003. The role of voice quality in communicating emotion, mood and attitude. *Speech Communication* 40/1–2. 189–212.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- Holmberg, Eva B. – Hillman, Robert E. – Perkell, Joseph S. – Guiod, Peter C. – Goldman, Susan L. 1995. Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures of female voice. *Journal of Speech and Hearing Research* 38/6. 1212–1223.
- Iseli, Markus – Alwan, Abeer 2004. An improved correction formula for the estimation of harmonic magnitudes and its application to open quotient estimation. In: *Proceedings of ICASSP*. Montreal, Quebec, Canada. 669–672.
- Kane, John – Drugman, Thomas – Gobl, Christer 2013. Improved automatic detection of creak. *Computer Speech & Language* 27/4. 1028–1047.
- Klatt, Dennis H. – Klatt, Laura C. 1990. Analysis, synthesis, and perception of voice quality variations among female and male talkers. *The Journal of the Acoustical Society of America* 87/2. 820–857.
- Malyska, Nicolas 2008. *Analysis of nonmodal glottal event patterns with application to automatic speaker recognition*. PhD-disszertáció. MIT, Massachusetts, USA.
- Markó Alexandra 2005. *A spontán beszéd néhány szuprasegmentális jellegzetessége. Monologikus és dialogikus szövegek összevetése, valamint a hümmögés vizsgálata*. PhD-disszertáció. ELTE, Budapest.
- Markó Alexandra 2011. A glottalizáció határjelző szerepe a felolvasásban. *Beszédkutató 2011*. 31–45.
- Markó Alexandra 2012a. Az irreguláris zöngé szerepe a magánhangzók határának jelölésében V(#)V kapcsolatokban. *Beszédkutató 2012*. 5–29.
- Markó, Alexandra 2012b. Boundary marking in Hungarian V(#)V clusters with special regard to the role of irregular phonation. *The Phonetician* 105–106. 7–26.

- Németh Géza – Olasz Gábor (szerk.) 2010. *A magyar beszéd. Beszédkutatás, beszédtechnológia, beszédinformációs rendszerek*. Akadémiai Kiadó, Budapest.
- Olasz Gábor 2013. Precíziós, párhuzamos magyar beszédadatbázis fejlesztése és szolgáltatásai. *Beszédkutatás 2013*. 261–270.
- Raitio, Tuomo – Kane, John – Drugman, Thomas – Gobl, Christer 2013. HMM-based synthesis of creaky voice. In *Proc. Interspeech 2013*. Lyon, France. 2316–2320.
- Redi, Laura – Shattuck-Hufnagel, Stefanie 2001. Variation in the realization of glottalization in normal speakers. *Journal of Phonetics* 29/4. 407–429.
- Ruinskiy, Dima – Lavner, Yizhar 2008. Stochastic models of pitch jitter and amplitude shimmer for voice modification. In: *IEEE 25th Convention of Electrical and Electronics Engineers in Israel*. Eilat, Israel. 489–493.
- Silén, Hanna – Helander, Elina – Nurminen, Jani – Gabbouj, Moncef 2009. Parametrization of vocal fry in HMM-based speech synthesis. In: *Proceedings of Interspeech 2009*. Brighton, UK. 1775–1778.
- Slifka, Janet 2007. Irregular phonation and its preferred role as a cue to silence in phonological systems. In *ICPhS 2007*. Saarbrücken, Germany. 229–232.
- Talkin, David 1995. A robust algorithm for pitch tracking (RAPT). In: Kleijn, W. Bastiaan – Paliwal, Kuldip K. (szerk.) *Speech coding and synthesis*. Elsevier, Amsterdam. 495–518.
- Tokuda, Keiichi – Kobayashi, Takao – Masuko, Takashi – Imai, Satoshi 1994. Mel-generalized cepstral analysis - A unified approach to speech spectral estimation. In: *Proceedings of ICSLP 1994*. Yokohama, Japan. 1043–1046.

A szerzők ezúton fejezik ki köszönetüket Olasz Gábornak a kézirattal kapcsolatos hasznos észrevételeiért. Emellett a szerzők köszönetet mondanak a percepció teszt résztvevőinek. A kutatást a PaeLife (Grant No. AAL-08-1-2011-0001) és az EIT-KIC (EITKIC_12-1-2012-001) projekt támogatta.

MAGYAR HANGSÚLYJELÖLÉSI SZÖVEGES ADATBÁZIS FEJLESZTÉSE ÉS REFERENCIA-VIZSGÁLATA

Olaszy Gábor – Abari Kálmán – Bartalis Mátyás

Bevezetés

Hangsúlyt is jelölő magyar mondatadatbázis eddig még nem állt a beszéd-kutatók rendelkezésére. Nemzetközi viszonylatban is kevés ilyen irányú kutatás található. A szón belüli hangsúly témakörével foglalkozik Goedemans–Hulst (2009) összefoglaló leíró jellegű munkája több nyelv vonatkozásában, összehasonlító nyelvészeti céllal. Angol mondatok szótagszintű hangsúly szintjeinek gépi megállapítására is vannak kísérletek, ezek főleg a nyelvtanulást, a kiejtés segítségét célozzák (Minematsu et al. 2002; Lu 2010). Konkrét hangsúlyadatbázisról, amely mérésekre, összehasonlításokra használható, nem találtunk szakirodalmi forrást, holott hangsúlycímkék a ma már rendelkezésre álló eszközökkel, manuális munkával támogatva bármely létező beszédadatbázishoz hozzáadhatók.

A magyar nyelv esetében a szón belüli hangsúlyozási szabály egyértelmű, alapvetően, ha hangsúlyos a szó, akkor az első szótagját hangsúlyozzuk. Ettől eltérő esetek is lehetnek (Kálmán–Nádasdy 2001), de ezekkel nem foglalkozunk. A mondaton belül pedig a szavak hangsúlyozása több tényezőtől is függ (kontextus, a beszélő akarata stb.). Sok esetben az értelmezés határozza meg, hogy melyik szóra kerül hangsúly a mondatban. Ezzel kapcsolatosan a leíró jellegű munkákon (É. Kiss et al. 1998; Kálmán–Nádasdy 2001; Varga 2002) túl már olyan kutatásokat is folytattak, amelyeknek algoritmusok kifejlesztése volt célja a magyar mondat szövegének automatikus elemzésére és a szavak hangsúly-kategóriákba való besorolására (Tamm–Olaszy 2005). Ezekre alapozva kezdtünk hozzá az első magyar szöveges hangsúlyjelölési mondatadatbázis elkészítéséhez, amely egységes ábrázolással biztosítja a mondat szavainak hangsúlycímkével történő ellátását kijelentő mondatokra (más modalitással nem foglalkoztunk). A modellünk szerint az elkészült szöveges adatbázisban minden mondat minden szava hangsúlycímkét kap. A szövegbe ágyazott hangsúlyjelöléseken túl előállítottuk a mondatok hangzó változatát is (gépi beszéddel), így nemcsak vizuális, hanem akusztikai formában is elérhetővé válik a kutató számára a teljes mondat hangsúlyképe. A munka három évet vett igénybe. Ez a hangsúlyadatbázis a későbbi kutatásokat szolgálja, egyrészt viszonyítási alap is lehet, másrészt konkrét, újszerű vizsgálatokra is ad inspirációt: hangsúlyjelölő algoritmusok tesztelése, beszédszintézis modellezése, gépi beszédfelismerés segítése, automatikus tarta-

lomkezelés támogatása, dialógusvizsgálatok szerkezeti kutatásának kiszélesítése, mondatelemzési kutatások tárgykörének bővítése, hangsúlykutatások támogatása stb. Az adatbázis tervezésekor az egyszerű, jól átlátható, egységes szerkezet kialakítása volt a cél. Ez – mint látni fogjuk – kompromisszumokat vont maga után. Az adatbázis a <http://magyarbeszed.tmit.bme.hu/hangsuly> címen az interneten is elérhető.

Anyag és módszer

Hangsúlyozzuk, hogy ebben a munkában a hangsúly fogalom nem azonos a fonológiai hangsúly fogalommal, a hangsúlyok meghatározása kizárólagosan akusztikus paramétereken és percepciók értékelésén alapul.

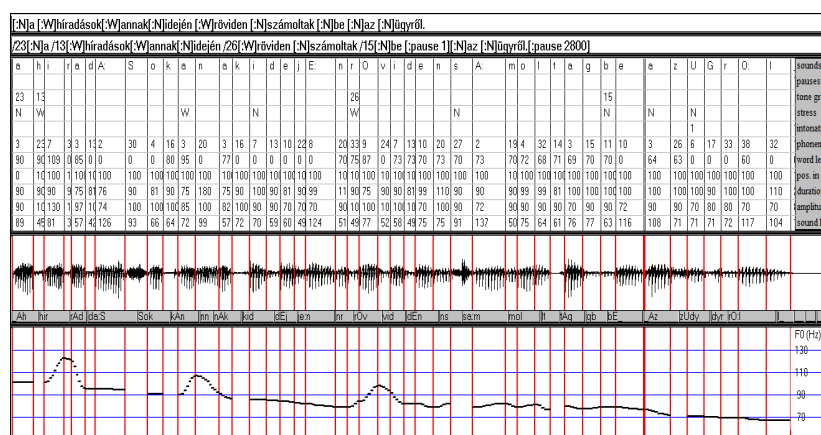
Az adatbázis nyelvi anyaga a BME TMIT-en korábban fejlesztett precíziós, párhuzamos beszédatadbázis 1866 db kijelentő mondatát tartalmazza. Ezek a mondatok irodalmi művekből származnak (Vicsi–Vígh 1998), tehát nem előre preparált szövegek, ezért a legkülönbébb mondat szerkezetek és mondat halmaz fordulnak elő. A mondatok nem alkotnak egységes szöveget, az adatbázis izolált mondatok halmazának tekinthető, mindenfajta tartalmi összefüggés nélkül. A kontextus mellőzése nagyfokú szabadságot biztosított, hogy a több lehetséges hangsúlycímké-kiosztás közül melyik kerüljön be az adatbázisba. A végleges döntés, mint később látni fogjuk, három magyar anyanyelvű beszélő nyelvi intuícióján alapult. A hangsúly jelölésére bináris elvet választottunk, kétféle állapotot különböztettünk meg: a szó vagy hangsúlyos (jelölése: H) vagy nem hangsúlyos (jelölése: -). (Itt jegyezzük meg, hogy e tanulmány szövegében több helyen a példamondatokban az egyszerűbb olvashatóság végett a hangsúlyos szavakat nem a H címkével jelöljük, hanem félkövér betűstílussal emeljük ki.) Nem foglalkoztunk a hangsúlyozás módjának és különböző elméleti fokozatainak finomabb részleteivel (mondathangsúly, fő- és mellékhangsúly, szakaszhangsúly stb.), mivel nincs kialakult metodológia ezen a téren. Célunk, hogy biztosítsuk az átláthatóságot és az egyszerű, egységes szerkezetet. További szempont volt, hogy minden mondatot csak egyfajta hangsúlycímké-kiosztással tároljunk el az adatbázisban.

A tervezett hangsúlycímkék meghatározásához gépi és kézi elemzést kombináltan használtunk. Alapként a Profivox gépi szövegfeldolvasó (Olaszy et al. 2000) hangsúlyjelölő moduljával vizsgáltattuk végig a mondatokat. Ez a program felszíni szövegelemzés és egyszerű szabályok segítségével állapítja meg a hangsúlyelosztást a mondaton belül. A program automatikusan felcímkézte a szavakat a kétféle hangsúly jellel. Egy felcímkézett mondat képe tehát a következő:

- (1) [-]a [H]híradások [H]annak [-]idején [H]röviden [-]számoltak [-]be [-]az [-]ügyről

Mivel a Profivox hangsúlyjelölő algoritmus csupán 70% körüli pontosságú, azaz olyan helyekre is tesz hangsúlyjelölést, amelyeknek hangsúlytalan-

nak kellene lenniük és fordítva, kétlépcsős ellenőrzést is lefolytattunk három magyar anyanyelvű személy bevonásával, hogy minél pontosabbá tegyük a mondatok szavainak a hangsúlyjeleket. Az első lépcsőben hangos olvasásos módszerrel kiejtettük a mondatokat, figyelembe véve az adott hangsúlyjeleket, és ha kellett, javítottuk a címkéken, csupán az anyanyelvi intuíción alapján (23%-ban kellett korrekciót végezni). Az esetleges hangsúlycímkejavításokat a többségi vélemény alapján fogadtuk el. Ezzel elkészült a hangsúlycímkekkel ellátott mondatok szöveges halmazának első verziója. A második ellenőrzéshez ezután minden mondatot beszédszintetizátorral hangzó formájává alakítottunk, melynek eredményeként a megadott hangsúlycímkek szerint állt elő a mondat kiejtett formája. A szintetizált beszéd alkalmazása biztosított egyfajta egységes beszédformát és hangzást (nem köthető adott személyhez a felolvasás). A hangsúlyok előállításánál kompromisszum volt, hogy csak az alaphangfrekvenciát (f_0 -t) változtattuk, intenzitás- és hangidőtartam-változtatást nem alkalmaztunk. Az f_0 -t megemeltük az első szótag magánhangzójában, majd fokozatosan visszavezértük a második szótag magánhangzójában. Az f_0 -csúcsot tehát szuperponáltuk a kijelentő mondat alap dallamvonulatára (1. ábra).



1. ábra

A szavakon elhelyezett hangsúlycímkek a vizsgált mondat szövegébe ágyazva (fent) és az ebből beszédszintézissel előállított hangzó forma adatmátrixa, hullámformája és alaphangfrekvencia-menete (legalul) a Profivox rendszerben (W = hangsúlyos)

A mondatok hangzó formáját ugyanazon három személy hallgatta meg. Most is elvégeztük az esetleges címkekorrekciókat a szövegben a többségi vélemény alapján. Mindösszesen 7%-ban kellett címkét javítani. A kétszeres

laboratóriumi ellenőrzés után előállt a végleges hangsúlycímkekkel ellátott szöveges adatbázis, amely alapvetően a három magyar ellenőrző személy döntéseit tükrözi.

Az adatbázis véglegesített hangsúlyjelöléseit a mondatok szövegében korrektnak tekintjük, amit úgy kell érteni, hogy nincs címkézési hiba, vagyis ahol hangsúlyt jelöltünk, ott a hangsúlyos ejtés nem okoz megértési zavart. Vannak olyan mondatok, amelyek többféle hangsúlykiosztással is ejthetők az értelmezés, illetve a közlési szándék szerint. Ezeknél a mondatoknál az egyik helyes formát tartalmazzák a jelölések.

A nyilvános teszt

Mivel célunk egy mindenki által biztos alapnak tekinthető, tudományos igényességgel megalkotott hangsúlyadatbázis elkészítése volt, felmerült a kérdés, hogy hogyan lehet módszertanilag is igazolni, hogy a szóhangsúlycímkek jó helyen vannak a mondatok szövegében. Véleményünk szerint, akkor tekinthetjük a fenti hangsúlyadatbázisban megadott hangsúlyjeleket validáltnak, ha sok ember véleménye alapján igazolni tudjuk, hogy a szavakhoz rendelt hangsúlyjelek megfelelő hangzást biztosítanak, nem okoznak zavart a prozódiai (esetleg értelmezési) megítélésben. Ezért az adatbázis referenciavizsgálatához webes nyilvános percepciók tesztet terveztünk.

A kidolgozott percepciók tesztünk a korszerű beszédminősítő rendszerekben megszokott összehasonlítási alapon működik (Fegyő 2010). Az összehasonlításos módszer alkalmazásához a következő gondolatmenet vezetett el: ha ugyanazon mondatot többféle hangsúlyozással, de – az f_0 -változtatást kivéve – teljesen azonos akusztikai tartalommal (ugyanazon hangidőtartamok, ugyanazon szünetek, azonos f_0 -alapgörbe) hozzuk létre, és ezeket páros összehasonlításban ütköztetjük, akkor az eredmények tükrözni fogják, hogy melyik a jobb hangsúlyozású. Azért vetettük el a mondatok abszolút minősítését (amely során egy mondat meghallgatását követően a minősítést végző személy egy kategóriába sorolja a mondatot), mert az ember számára nehezebb jó/rossz hangsúlyozású, vagy akár többértékű kategóriákba sorolni a mondatokat. Az egyszerűbb döntéshozatal mellett a páros összehasonlításnak másik előnye, hogy a döntési helyzetben szereplő, géppel előállított mondatok csak a hangsúlyparaméterben különböznek, minden más változó hatását ki tudjuk küszöbölni.

Ennek megfelelően az adatbázis ellenőrzött és jóváhagyott hangsúlycímkeit tartalmazó (továbbiakban A-típusú) mondataihoz készítettünk további két preparált szintetizált mondatot is, amelyek csak a hangsúlycímkekben tértek el az A-típusú mondatoktól (ezek a B- és C-típusú mondatok). Ezeket is manuálisan, a Profivox fejlesztői rendszerével (Olaszy et al. 2000) szintetizáltuk. Így hangsúlycímkekkel ellátott mondatok és hozzájuk kapcsolódó hangálmányok három halmaza állt elő:

A-típusú mondatok: mondat a helyesnek tartott hangsúlyozással (ezek a mondatok alkotják a hangsúlyadatbázist).

B-típusú mondatok: ugyanazon mondat, melynek minden szava hangsúlytalan címkét kapott (a hangsúlyt levettük, a dallam vonulata maradt).

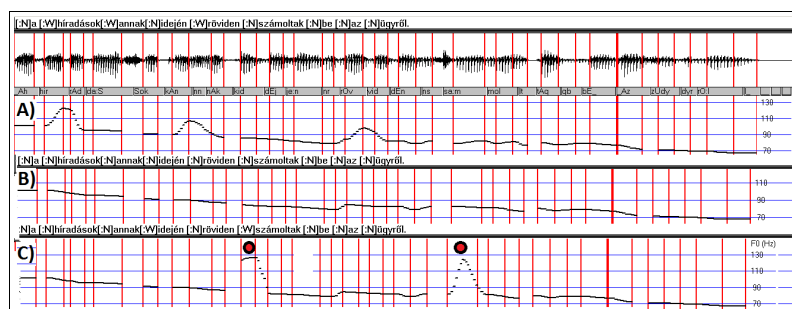
C-típusú mondatok: Kontrasztot próbáltunk létrehozni az A-típusú mondatokkal szemben, azaz ugyanazon mondatot megpróbáltuk rossz hangsúlyozással megvalósítani (a hangsúlyt többnyire egy korábban nem hangsúlyozott szóra tettük át, a dallam vonulata változatlan maradt). Itt megjegyezzük, hogy ennek a változatnak az elkészítése jelentette a legnagyobb nehézséget, mivel bizonyos mondatok tartalmához nem lehetett az A-típusú mondatról jól elkülöníthető rossz hangsúlyozást készíteni.

Az 1. táblázatban példát láthatunk, hogy a kiinduló A-típusú alakból hogyan származtattuk a B- és C-típusú mondatokat.

1. táblázat: Egy példa a B- és C-típusú mondatok származtatására

A-típus: jó hangsúlyozás	[-]a [H]híradások [H]annak [-]idején [H]röviden [-]számoltak [-]be [-]az [-]ügyről
B-típus: neutrális forma	[-]a [-]híradások [-]annak [-]idején [-]röviden [-]számoltak [-]be [-]az [-]ügyről
C-típus: rossz hangsúlyozás	[-]a [-]híradások [-]annak [H]idején [-]röviden [H]számoltak [-]be [-]az [-]ügyről

A háromféle mondat hangzó formáiban lévő különbségekre a 2. ábrán láthatunk példát.



2. ábra

A különböző hangsúlyozással előállított mondatok a szövegbe ágyazott címkéikkel és a megvalósított három alaphangfrekvencia-menet a Profivox rendszerben (W = hangsúlyos)

A nyilvános teszthez tehát $1866 \times 3 = 5598$ db szintetizált mondatot állítottunk elő. Minden mondat három különböző hangsúlymintázattal készült el. Az A-típusú mondatokat referenciamondatoknak tekintettük. A tesztelők számára a páros összehasonlításokhoz AB és AC mondatpárokat hoztunk létre: (i) AB

mondatpár: az A-típusú mondatot ütköztettük a B-típusú megfelelőjével, tehát az elhangzási sorrendet tekintve véletlenszerűen vagy A-B vagy B-A mondatpárt hallott a tesztelő. (ii) AC mondatpár: Az A-típusú mondatot ütköztettük a C-típusú megfelelőjével. Itt is véletlenszerű sorrendben A-C vagy C-A mondatpárt hallott a tesztelő.

Így minden A-típusú mondat összehasonlításra került a saját B- és C-típusú változatával a teszt során. Mivel sok mondatot kellett ilyen módon összehasonlítani, sok tesztelőre volt szükség ahhoz, hogy statisztikailag értékelhetőek legyenek a válaszok. Több ezer tesztelőnek küldtük ki a felhívást a teszt elvégzésére, zömében egyetemi hallgatóknak és dolgozóknak, szakmai közösségeknek az ország minden tájára. Minden tesztelőtől általános adatokat is kértünk: az életkorát, nemét és egy városnevet, ahol él, valamint azt, hogy hangszórón vagy fejhallgatón hallgatta-e a mondatokat. Egy-egy tesztelő 40 mondatpárt hallgatott meg, 20 db AB és 20 db AC mondatpárt, és mindegyikről ítéletet mondott. Egy tesztfolyamat átlagosan 15 percet vett igénybe. Minden tesztelő más-más 40 mondatpárt hallott. Az éppen hallott mondatpárt akárhányszor lejátszhatta a tesztelő, mielőtt meghozta az ítéletet. Egy-egy tesztelő többször is elvégezhetette a tesztet, ha újból bejelentkezett, minden alkalommal más-más 40 mondatos blokkot hallott.

A tesztben feltett kérdés és a rá adható négy válasz (melyből egyet kellett bejelölnie a tesztalanyoknak) a következő volt:

„Kérjük, ítélje meg az elhangzott mondatok hangsúlyozását.

1. Az első hangsúlyozása jobb, természetesebb, mint a másodiké.
2. A második hangsúlyozása jobb, természetesebb, mint az elsőé.
3. Mindkét mondat hangsúlyozása jó.
4. Egyik mondat hangsúlyozása sem jó.”

A kérdés és a rá adható válaszok megfogalmazásánál ügyeltünk arra, hogy a szubjektív beszédminősítés páros összehasonlítások során alkalmazott szokásos értékelése elvégezhető legyen. A teszt eredményeinek kiértékeléséhez a beszédtechnológiában elterjedt CMOS (Comparison Mean Opinion Score) (ITU-T 1996: 23–25) alapú értékelés egyszerűbb változatát használtuk. A tesztre adott 4 válasz egy 3 fokozatú értékelési skála kialakítására adott lehetőséget. Ha a meghallgatás során a tesztalanyok az A-típusú mondatot ítélték jobb hangsúlyozásúnak, akkor 1 pontot ért a válasz, ha a B- vagy C-típusú mondatot, akkor –1 pontot. A 3. és 4. válasz esetén 0 pontot kapott a válasz.

Az értékelés során a pontszámok átlagát, a CMOS-értékeket használtuk fel. Egy példán keresztül mutatjuk be, hogy egyetlen mondatra vonatkoztatva a pontszámok átlaga milyen jelentéssel bír (2. táblázat). A példamondat: *Ne rólam, hanem a **jelenségről** beszéljünk.*

2. táblázat: Az adatbázis egyetlen mondatának CMOS-értékei 11 tesztelő értékelésében (zárójelben a tesztelők száma található)

CMOS	CMOS (AB mondatpár)	CMOS (AC mondatpár)
0,91 (N = 11)	0,8 (N = 5)	1 (N = 6)

A 2. táblázat szerint 5 tesztelő AB-típusú mondatpárt, 6 tesztelő pedig AC-típusút hallgatott meg. Az AB mondatpár oszlop 0,8-as értéke azt jelenti, hogy az 5 tesztelő többsége az A-típust jelölte meg az összehasonlításból jobbnak (ha mindannyian az A-t jelölték volna meg, akkor 1 lenne az érték). Négy tesztalany az A-típusú mondatot részesítette előnyben, egy pedig egyformának ítélte a B-típusú mondattal. Azaz $0,8 = (1 + 1 + 1 + 1 + 0)/5$. Az ugyanezen AC mondatpár oszlopában az 1-es érték azt jelenti, hogy mind a hat tesztalany az A-típusú mondatot preferálta. A 2. táblázat CMOS oszlopában az összesített, az AB és AC mondatokra egyaránt vonatkozó ítéletek átlaga szerepel (0,91).

A CMOS-értékek statisztikai elemzésére egy- és kétmintás t -próbát, valamint kétszemponos varianciaelemzést használunk. A válaszkategóriák eloszlását gyakorisági táblázatokkal írjuk le, és χ^2 -próbával elemezzük.

Az elkészült hangsúlyadatbázis szerkezete és jellemzői

Az adatbázis 1866 kijelentő mondata a megadott hangsúlycímkekkel egyetlen TXT-fájlban található, és minden mondat saját azonosítóval van el látva (3. ábra). A mondatok szintetizált változatát WAV formátumban tároljuk 22 kHz-es mintavételi frekvenciával és 16 bites kvantálással ugyanazon azonosítóval, mint ami a szöveges mondathoz tartozik.

[H]apámnak [H]negyvenöt [-]tavaszára [H]sikertült [H]visszasietnie
 [H]erőt, [H]egészséget, [H]szolgáljuk [-]a [H]rendőrséget
 [H]feltétlenül [-]elgőzösítik
 [H]megcsodálhattuk [H]netanjahu [-]és [H]arafat [H]partnerszellemét
 [-]pillanatnyilag [H]megfeledezett [H]bertold [H]árnyalakjáról
 [H]benjamin [H]netanjahu [H]arnold [H]kollerrel [-]találkozott
 [H]jizzadságkészlete [H]kimeríthetetlennek [-]látszott
 [H]nem [-]jó [H]mindent [H]megrohasztani, [H]meggyengíteni, [H]aláaknázni
 [H]mikszáthnak [H]kezdőként [H]nem [-]volt [H]semmije

3. ábra

Mondatpéldák a hangsúlyjelölési adatbázisból

A hangsúlyadatbázis mondatait jellemző szavak száma és hangsúlyok száma változók keresztábláját a 3. táblázat tartalmazza.

3. táblázat: A hangsúlyok számának és a szavak számának kétdimenziós gyakorisági táblázata az adatbázis 1866 mondatában

	Hangsúlyok száma								
	1	2	3	4	5	6	7	8	Σ
Szavak száma	2	1	0	0	0	0	0	0	1
	3	3	5	0	0	0	0	0	8
	4	20	23	6	4	0	0	0	53
	5	54	98	32	5	0	0	0	189
	6	60	155	80	19	2	0	0	316
	7	56	161	145	35	7	1	0	405
	8	37	145	119	54	9	0	0	364
	9	14	83	98	51	5	1	0	252
	10	4	38	60	45	3	0	3	153
	11	1	8	21	25	10	0	0	66
	12	0	7	12	11	7	0	0	37
	13	0	1	6	4	6	0	1	18
	14	0	0	1	2	0	1	0	4
	Σ	250	724	580	255	49	3	4	1866

A 3. táblázatból kiolvasható, hogy a legtöbb mondat 6–8 szót tartalmaz, és a kis vagy nagy szószámok esetén egyre csökken a hangsúlyos szavak gyakorisága. Az adatbázisban 2–3 hangsúlyszám a leggyakoribb, azonos mértékben tartalmaznak 1 és 4 hangsúlyt a mondatok. A hangsúlyszám növekedésével, egészen a maximális 8 hangsúlyig a mondatok gyakorisága jelentősen csökken. Mindösszesen 4 mondat van, amelyekben minden szó hangsúlyjellel van ellátva, például: *Minden lecsillapodott, elsimult, elcsitult.* A legtöbb hangsúlyjel 8 darab, ez egyetlen mondatban fordul elő: *Bejön a mama, foga nincs, angolul nem tud, de szépen mosolyog.* Hét hangsúlyjel 4 mondatban, hat 3-ban, öt 49-ben fordul elő. Mindössze egyetlen hangsúlyos szó (a többi hangsúlytalan) 250 mondatban található. Meg is fordíthatjuk a vizsgálatot, és a [-] címkéket is vizsgálhatjuk. Egyetlen hangsúlytalan jelölés 20 mondatban van, kettő már 85-ben. Azon mondatok száma, amelyekben több a hangsúlyos szó, mint a hangsúlytalan: 136. Mindezek a számadatok azt mutatják, hogy számos szempont szerint kaphatunk adatokat az adatbázis szerkezeti elemeiről. A hangsúlyadatbázis sok szempontú vizsgálatára webes lekérdező-felületet fejlesztettünk (Abari–Olaszy 2014). A fenti adatok ezen alkalmazás használatával is kinyerhetők.

Hangsúlymintázatok a mondathossz függvényében

A mondat szavaira tett hangsúlyjelek sorozata adja a hangsúlymintázatot. A hangsúlymintázat annyi jelet tartalmaz, ahány szó van a mondatban. Az

egyetlen kétszavas mondat hangsúlymintája H-, ami azt jelenti, hogy az első szó hangsúlyos, a második hangsúlytalan. A mondat: **Feltétlenül** előzősítik.

A három és néhány többszavas mondatban előforduló hangsúlymintázatok a 4. táblázatban láthatók.

4. táblázat: Hangsúlymintázatok a 3–6 szavas mondatokban. A számok az adatbázisbeli előfordulások számát jelzik. A zárójelben lévő százalékok az adott szószámú mondatokon belüli arányt fejezik ki

3 szavas mondatok			
H-- 3 (37,5%)	HH- 3 (37,5%)	-HH 1 (12,5%)	H-H 1 (12,5%)
4 szavas mondatok (csak a 10% feletti gyakoriságúak)			
H--- 13 (24,5%)	HH-- 8 (15,1%)	-H-- 6 (11,3%)	-HH- 6 (11,3%)
5 szavas mondatok (az első 4 legnagyobb gyakoriságú)			
H---- 26 (13,8%)	-H--- 22 (11,6%)	-HH-- 22 (11,6%)	-H-H- 18 (9,5%)
6 szavas mondatok (az első 4 legnagyobb gyakoriságú)			
-H-H- 36 (11,4%)	H-H-- 24 (8,0%)	-H---- 22 (7,0%)	H----- 18 (6,0%)

A hangsúlymintázat vizsgálatával a szövegtartalom és a hangsúlyozás közötti összefüggéseket lehet vizsgálni, például hogy az azonos mintázathoz tartozó szövegekben van-e mondat szerkezeti hasonlóság.

A nyilvános percepciók teszt eredményei

Hipotéziseink a páros összehasonlításokra vonatkozó válaszokkal kapcsolatosan a következők voltak.

1. hipotézis: Az A-típusú mondatokat (hiszen ez képviseli a korrekt hangsúlyozást) többen választják jó hangsúlyozásúnak, mint a B- és C-típusúakat.

2. hipotézis: Az AB mondatpárok esetén gyakrabban kapunk 3. választ, mint akkor, ha a tesztalany AC mondatpárt hall. Ezt a hipotézist arra alapoztuk, hogy az A- és B-típusú mondatok között kisebb az akusztikai különbség, mint az A- és C-típusúak között (a C-típusú mondatokban elrontott hangsúlyozást igyekeztünk megvalósítani). Azoknál az AB mondatpároknál, amelyikek A-típusú mondatában mindössze egy szón van hangsúly, az A- és B-típus közötti akusztikai különbség szinte alig hallható. Ezért nagy valószínűséggel kaphatjuk a 3. választ is (egyformán jó mindkét mondat hangsúlyozása).

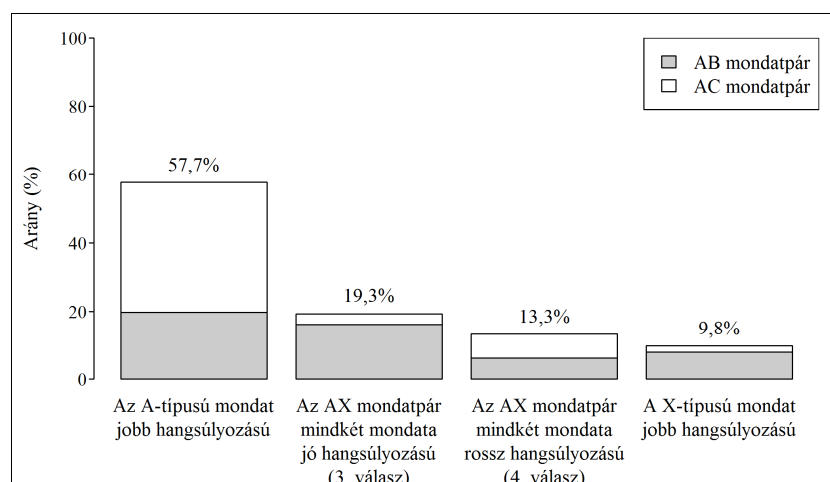
3. hipotézis: Ha a tesztelő az AC mondatpárt hallja, akkor gyakrabban fog az A-típusú mondatok javára dönteni, mint az AB mondatpároknál. Ezt a hipotézist arra alapozzuk, hogy az A- és C-típusúak között nagyobb az akusztikai különbség, mint az A- és B-típusúak között.

A nyilvános tesztet 351 személy (188 férfi és 163 nő) végezte el 2013 október–november hónapjaiban. A tesztalanyok átlagéletkora 30 év, szórása 14 év. A legfiatalabb tesztkitöltő 17, a legidősebb 85 éves volt. A kitöltők kétharmada 28 évesnél fiatalabb volt. A tesztalanyok 60%-a budapesti, 15%-a debreceni lakos, a többi egyéb településekről végezte a tesztet. A teszteredmények közreadását a válaszkategóriák gyakorisági elemzésével kezdjük, majd rátérünk a CMOS-értékek vizsgálatára.

Megvizsgáltuk, hogy a 351 válaszadó a páros összehasonlítások során a 4 lehetséges választ milyen arányban választotta. Az AB és AC mondatpárookra AX mondatpárként hivatkozunk. Az X jelölés tehát az A-típusú mondatnál párban hallott B- vagy C-típusú megfelelőt jelenti. A teszt során kiválasztható 4 válasz így egyértelműen besorolható a következő 4 kategória valamelyikébe: (i) az A-típusú mondat a jobb hangsúlyozású, (ii) az X-típusú mondat a jobb hangsúlyozású, (iii) mindkét mondat (A- és X-típusú is) jó hangsúlyozású és (iv) mindkét mondat (A- és X-típusú is) rossz hangsúlyozású. A $351 \times 40 = 14\,040$ db válasz alapján az egyes kategóriák arányát a 4. ábra tartalmazza. A 4. ábráról leolvasható adatok mindhárom hipotézis beigazolódását megmutatják.

Az első hipotézis teljesülését mutatja az a tény, hogy az összes válasz 57,7%-ában az A-típusú mondatokat ítélték jobb hangsúlyozásúnak a vele párban lévő B- vagy C-típusú mondatnál szemben, míg ellenkező válasz csak az esetek 9,8%-ában született. Nagyon leegyszerűsítve azt mondhatjuk, hogy mind a négy válaszlehetőséget figyelembe véve, minden második válasz az A-típusú mondatot jelölte jobb hangsúlyozásúnak, és csak minden tizedik az X-típusút. A kimaradt esetek a 3. és 4. válasz között oszlanak el. Az AX mondatpár mindkét tagját az esetek 19,3%-ban jó hangsúlyozásúnak, 13,3%-ban pedig rossz hangsúlyozásúnak ítélték a tesztalanyok. A válaszkategóriák eloszlása nem véletlenszerű, χ^2 -próba szerint a hangsúlycímke-módosítás (az A-, B- és C-típusú mondatok kialakításának) hatása szignifikáns volt ($\chi^2 = 8244,4$, $df = 3$, $p < 0,001$) a válaszokra.

Tovább erősíthetjük az első hipotézis igazolását, ha a 4. ábra két bal oldali oszlopának részarányát összegezzük. Így azt kapjuk, hogy az A-típusú mondatokat 77%-ban nem ítélték rossz hangsúlyozásúnak a tesztelők. E kijelentés mögött a válaszkategóriák egyszerű abszolút értékelésének lehetősége húzódik meg. Ugyanis a páros összehasonlítások során implicit módon az A-típusú mondatokról abszolút döntést hoztak a tesztalanyok a következő elv alapján: az A-típusú mondat nem képvisel rossz hangsúlyozású mondatot, ha a tesztalany a 4. ábra két bal oldali oszlopában leírt kategóriából valamelyiket választotta, viszont rossz hangsúlyozású, ha a jobb oldali két kategóriát választotta. Vagyis az A-típusú mondat jó hangsúlyozású, ha jobb hangsúlyozású az X-típusú megfelelőjénél (4. ábra bal első oszlop), vagy az AX mondatpárban jó hangsúlyozásúnak ítélték (4. ábra második oszlop balról).



4. ábra

A nyilvános teszt négy válaszáinak eloszlása. A válaszok százalékos gyakoriságát megjelenítő oszlopokban külön jelöltük az AB (sötét) és AC (világos) mondatok arányát

Az ábráról leolvasható, hogy az egyes válaszkategóriákban az AB és AC mondatpárok aránya eltér (ez az eltérés is szignifikáns: $\chi^2 = 2592,8$, $df = 3$, $p < 0,001$). A második hipotézis a 3. válasz egyenetlen eloszlását feltételezi az AB és AC mondatpárok között. Ez teljesül, mivel az AX mondatpár mindkét tagját jónak minősítő válaszokban az AB és AC mondatpárok között az arány 83,7-16,3%, tehát csak minden ötödik ilyen típusú válasz származik AC mondatpárból (a 4. ábra második oszlopa). A harmadik hipotézis az A-típusú mondatokat jobbnak minősítő válaszokban (a 4. ábra első oszlopa) jósol egyenetlen eloszlást, de fordított előjellel. Az AB és AC mondatpárok aránya 34,2%-65,5%, azaz az A-típusú mondatokat előnyben részesítő válaszok kétharmadát AC mondatpár meghallgatása során kaptuk. Ez bizonyítja a harmadik hipotézis teljesülését is. A 4. ábra maradék két oszlopára is elvégezhetjük a fenti elemzéseket. Az AX mondatpár mindkét tagját egyaránt rossznak minősítő válaszok (harmadik oszlop) közel azonos arányban származnak AB és AC mondatokból (46,3%-53,7%). Az X-típusú mondatokat jobb hangsúlyozásúnak minősítő válaszok (negyedik oszlop) túlnyomó többsége, 81,5%-a AB típusú mondatpárból származik, csak a maradék 18,5% jön AC mondatpárokból.

A válaszok értékelését a CMOS-értékek alapján is elvégeztük. A továbbiakban ezt a mérőszámot arra használjuk, hogy különböző csoportokban az A-típusú mondatokra vonatkozó kedvező döntések számát egymással összehasonlítsuk.

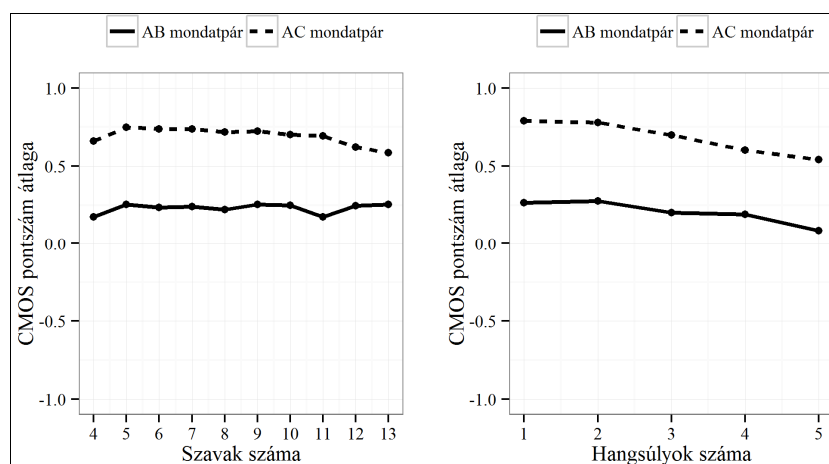
Kiszámoltuk az egész adatbázisra vonatkozó átlagos CMOS-értéket az összes AB és összes AC mondatpár esetén (5. táblázat). A pozitív értékek mindkét mondatpár esetén megmutatják, hogy szívesebben választják a tesztalanyok az A-típusú hangsúlyozással ellátott mondatokat. Az AC mondatpárok esetében ez sokkal hangsúlyosabb, mivel az akusztikai különbség nagyobb, mint az AB mondatpároknál. Mindkét várható érték szignifikánsan eltér a 0-tól (egymintás t-próbával, AB mondatpár esetében $t = 23,7$, $df = 1856$, $p < 0,001$; AC mondatpár esetében $t = 96,4$, $df = 1856$, $p < 0,001$;). Az AB és AC mondatpárookra vonatkozó eltérések is szignifikánsan eltérnek egymástól (kétmintás t-próbával, $t = -40,2$, $df = 3711$, $p < 0,001$). A CMOS-értékeken alapuló fenti három statisztikai próba eredménye egyben az első és harmadik hipotézis teljesülését is jelenti.

5. táblázat: Az átlagos CMOS-értékek az összes mondatpárra

	Átlag	Szórás	N
AB mondatpár	0,24	0,71	7020
AC mondatpár	0,72	0,52	7020

Az átlagos CMOS-értékek változását vizsgálhatjuk a tesztmondatok különböző jellemzői szerint is. A szavak száma és a hangsúlyok száma szerinti átlagos CMOS-pontszámok változását az 5. ábra tartalmazza. A 2, 3 és 14 szószámmal rendelkező mondatok a kis mintaelemszám miatt kimaradtak a vizsgálatból. Hasonló megfontolások miatt a hangsúlyszámok szerinti csoportosításban nem szerepelnek a 6–8 hangsúllyal rendelkező mondatok.

Az 5. ábra bal oldali része megmutatja, hogy a szavak száma nem befolyásolja az ítéleteket sem az AB, sem az AC mondatpárok esetén. Két szempontos varianciaelemzés segítségével egyedül a mondatpárok hatása bizonyult szignifikánsnak: $F(1, 13899) = 2135,6$; $p < 0,001$. Az 5. ábra jobb oldali részén a hangsúlyszám emelkedésével csökkenés figyelhető meg az A-típusú mondatokra vonatkozó kedvező döntések számában. Szignifikáns hatása van a mondatpár változónak: $F(1, 13915) = 2157,7$; $p < 0,001$, valamint a hangsúlyok számának is: $F(1, 13915) = 114,1$; $p < 0,001$. A CMOS-érték csökkenésének oka a hangsúlyszám növekedésével további, részletesebb vizsgálatokat kíván.



5. ábra

A CMOS-átlagok változása a szavak számának és a hangsúlyok számának függvényében

Vizsgálhatjuk az egyes mondatokra vonatkozó egyedi CMOS-értékeket is a válaszok függvényében. Ezt akár minden mondatra el lehet végezni. Itt most csak néhány érdekes eset említünk. Például megkerestük a szélsőséges CMOS-értékekkel rendelkező mondatokat. Az elemzést itt is külön végeztük az AB és AC mondatpárookra, és csak azokat a mondatokat vizsgáltuk, amelyekre legalább 5-5 válasz érkezett.

Az A-típusú mondatok számára legkedvezőtlenebb eset a -1 -es átlagos CMOS-pontszám, amely csak egyetlen mondatnál, az AB mondatpár tesztelési válaszaiból adódik. Ez azt jelenti, hogy minden tesztelő a neutrális hangsúlyozását tekintette jobbnak nem pedig az A-típusút:

(2) [-]a [H]fehér [-]köpenyes [-]ember [H]meg [-]se [H]fordult.

Az eredmény például magyarázható azzal, hogy a mondat tartalma nem kíván különösebben hangsúlyozást, anélkül sem sérül az értelmezése.

Az AC mondatpárok vizsgálatában a legkevésbé kedvező ítélet CMOS-értéke $-0,66$, ami egyetlen mondatpárra vonatkozik. Itt majdnem minden tesztelő a C-típusú mondatot tartotta jobbnak. Az A és C-típusú mondatok a következők:

- (3) A) [H]arról, [-]ami [H]kísérletileg [H]nem [-]közelíthető [-]meg, [H]hallgat.
 C) [H]arról, [-]ami [H]kísérletileg [-]nem [H]közelíthető [H]meg, [H]hallgat.

A két hangsúlyozás összehasonlításából látható, hogy az adatbázist fejlesztőknek nem sikerült megvalósítani eléggé karakteresen a rossz hangsúlyozást a C-típusú (alsó) mondatban, hiszen három esetben a hangsúlyok ugyanazon a szavakon vannak mindkét mondatban. A mondatok közötti akusztikai különbség tehát kicsi. A tesztelők ítéletei ezt jelzik. Szerencsére csak egy ilyen mondat van a rendszerben. Mutatunk néhány példát olyan AB mondatpárokból is, amelyekben minden tesztelő az A-típusú mondatot tartotta jobbnak (CMOS = 1).

- (4) A) [-]a [H]hogyan [-]már [H]megértette, [-]a [H]miért [-]még [-]nem.
 A) [H]nagyon [-]optimista [-]vagyok [-]a [-]jövővel [-]kapcsolatban
 A) [-]az [H]ügyvéd [-]a [H]pénzéért [-]a [H]feladatokat [-]is [-]kijelöli

A fentiek folytatásaként, ilyen részletes vizsgálatok elvégzése után, akár korrigálhatjuk is a címkéket a hangsúlyadatbázis néhány mondatában, de ezek a korrekciók a mondatok többségének megítélését az eddigi adatok szerint nem befolyásolják. Ilyen kutatást a jövőben terveznek elvégezni a cikk szerzői, de bárki előtt szabad az út a kérdéskör tetszőleges irányú vizsgálatára.

Összefoglalás

Az első magyar hangsúlyadatbázist mutattuk be, olyan szöveges mondatgyűjteményt, ahol a mondat minden szava hangsúlycímkével van ellátva. Nyilvános percepciós teszttel igazoltuk, hogy a megadott hangsúlycímkék nem képviselnek rossz hangsúlyozást. Így az adatbázis validáltnak tekinthető az adott mondatok vonatkozásában. Az adatbázis önmagában is számos kutatási kérdés vizsgálatára ad lehetőséget. Ezekből néhányat felvillantottunk. A nagyszámú tesztelő válaszainak részletesebb, több szempontú vizsgálata is lehetséges, erre egy külön tanulmányban térünk vissza. Tovább bővíti a kutatási teret az a tény is, hogy a hangsúlyadatbázis minden mondata 12 személy ejtésében is rendelkezésre áll a TMIT adatbázisában (Olaszy 2013), tehát összehasonlíthatók a természetes ejtések a hangsúlyadatbázis adataival. Ezzel a valós ejtés és az adatbázis közötti kapcsolat is vizsgálható. A webes lekérdezőfelülettel kiegészített adatbázis elérhető az interneten is (Abari–Olaszy 2014).

Irodalom

- Abari Kálmán – Olaszy Gábor 2014. Magyar hangsúlyadatbázis az interneten kutatáshoz, oktatáshoz. In Tanács Attila – Varga Viktor – Vincze Veronika (szerk.): *X. Magyar Számítógépes Nyelvészeti Konferencia*. Szegedi Tudományegyetem, Informatikai Tanszékcsoport, Szeged, 347–356.
- É. Kiss Katalin – Kiefer Ferenc – Siptár Péter 1998. *Új magyar nyelvtan*. Osiris Kiadó, Budapest.
- Fegyő Tibor 2010. Gépi beszédminősítés távközlési rendszerekben. In: Németh Géza – Olaszy Gábor (szerk.): *A magyar beszéd*. Akadémiai kiadó, Budapest. 547–555.

- Goedemans, Rob – Hulst, Harry 2009. StressTyp: A database for word accentual patterns in the World's languages. In: Everaert, Martin – Musgrave, Simon – Dimitriadis, Alexis (eds.): *The use of databases in cross-linguistic studies*. Mouton de Gruyter, Berlin–New York. 235–282.
- ITU-T: P.800. 1996. *Methods for subjective determination of transmission quality*. <http://ip66.twbbs.org/VoIP/Measurement/p.800.pdf>.
- Lu, Jingli 2010. *CASTLE: A computer-assisted sentence stress teaching and learning environment*. PhD thesis. Massey University, Manawatu, New Zealand.
- Minematsu, Nobuaki – Kobashikawa, Satoshi – Hirose, Keikichi – Erickson, Donna 2002. Acoustic modeling of sentence stress using differential features between syllables for English rhythm learning system development. In Hansen, John H. L. – Pellom, Bryan L. (eds.): *Proceedings of Interspeech*, ISCA. 745–748.
- Kálmán László – Nádasdy Ádám 2001. A hangsúly. In Kiefer Ferenc (szerk.): *Strukturális magyar nyelvtan 2. Fonológia*. Akadémiai Kiadó, Budapest. 393–467.
- Olaszy, Gábor – Németh, Géza – Kiss, Géza 2001. Hungarian audiovisual prosody composer and TTS development tool. In Puppel, Stanislaw – Grazina, Demenko (eds.): *Prosody*. Poznan. 167–178.
- Olaszy Gábor 2013. Precíziós, párhuzamos magyar beszédatadátbázis fejlesztése és szolgáltatásai. *Beszédkutatás* 2013. 261–270.
- Olaszy Gábor – Németh Géza – Olaszi Péter – Kiss Géza – Zainkó Csaba – Gordos Géza 2000. Profivox – a Hungarian TTS system for telecommunications applications. *International Journal of Speech Technology* 3–4. 201–215.
- Tamm Anne – Olaszy Gábor 2005. Kísérlet automatizált szövegelemzési módszerek kialakítására a szóhangsúlyok meghatározásához. In Alexin Zoltán – Csendes Dóra (szerk.): *III. Magyar Számítógépes Nyelvészeti Konferencia*. Szegedi Tudományegyetem Informatikai Tanszékcsoport, Szeged. 383–393.
- Varga László 2002. *Intonation and Stress: Evidence from Hungarian*. Palgrave Macmillan, New York.
- Vicsi Klára – Vigh Attila 1998. Az első magyar nyelvű beszédatadátbázis. *Beszédkutatás* '98. 163–177.

A kutatást a Paelife (Grant No. AAl-08-01-2011-0001) és az EITKIC_12-1-2012-001 projektek támogatták.

TEMPÓVÁLTOZÁSOK A VIZSGÁLT SZAKASZ NAGYSÁGÁNAK FÜGGVÉNYÉBEN

Kohári Anna

Bevezetés

A hangzók időtartamát számtalan tényező egyszerre befolyásolja a beszéd-produkció során, úgymint a hangzó minősége, környezete, a hangsorban elfoglalt helye, hangsúlyos vagy hangsúlytalan volta, a szó hosszúsága, a szó helyzete a közlésben, a beszédtempó, a beszédmód stb. (l. van Santen 1992; Gósy 2004; Gósy–Beke 2010). A folyamatos beszédben a különböző hangzók időtartama éppen ezért nehezen jósolható meg. Az időtartamok változosságának leírására és modellezésére többféle megközelítés is született.

A beszédritmusmodellek egyike kifejezetten a folyamatos beszéd időtartambeli sajátosságait próbálja megragadni (Ramus et al. 1999; Grabe–Low 2002; Dellwo 2009). A modellben az ún. mássalhangzós szakaszt és a magánhangzós szakaszt definiálták a hangzó beszéd két alapegységeként. Mássalhangzós szakasz alatt az egymást közvetlenül követő mássalhangzókat értjük, amelyeket nem választ szét magánhangzó. Így a hangsorok felépítésétől függően egy mássalhangzós szakasz állhat egy, kettő vagy akár több mássalhangzóból is. A módszer nem veszi figyelembe a néma szüneteket és hezitálásokat, kizárólag a beszédhangok mért időtartamait. A szakaszok átívelhetnek a szóhatárokon, mondathatárokon és a szüneteken is (Ramus et al. 1999; Grabe–Low 2002). A mássalhangzós szakasz időtartamát pedig a benne lévő hangzók időtartamának összege adja. A *teáskanál* szóban például a [ʃ] és a [k] együttesen alkot egy mássalhangzós szakaszt. E két hang időtartamának összege az adott mássalhangzós szakasz időtartama is egyben. A példában a [t], a [n] és a [l] hangzók pedig önmagukban alkotnak egy-egy mássalhangzós szakaszt a definíció értelmében. A mássalhangzós szakaszhoz hasonló módon definiálható a magánhangzós szakasz is. Az egymást közvetlenül követő egy, két vagy több magánhangzó, amelyet nem választ szét mássalhangzó, alkot egy magánhangzós szakaszt. A magánhangzós szakaszok és mássalhangzós szakaszok időtartamának segítségével különböző mérőszámokat hoztak létre az egyes nyelvek sajátosságainak a mérésére (Ramus et al. 1999; Grabe–Low 2002). A mérőszámok egy része a magánhangzós szakaszok és a mássalhangzós szakaszok időtartamának arányain, szórásán alapul (Ramus et al. 1999). Egy másik típus (Pairwise Variability Index, PVI) azonban jelentőséget tulajdonít a magánhangzós és mássalhangzós szakaszok sorrendjének is. Ez a mérőszám típus nem a különböző szakaszok tiszta időtartamából indul ki, hanem az egymást követő egységek időtartamainak különbségét veszi ala-

pul (Grabe–Low 2002), az időtartamokat tehát minden esetben viszonyítja a környező egységek időtartamához.

Az időtartamokon alapuló beszédritmus-mérőszámok jelentőségét mutatja, hogy az általuk kapott eredmények szoros összefüggést mutattak egy korábbi, a percepció tapasztalatból kiinduló elmélet beszédritmusosztályaival. A múlt század közepén három beszédritmusosztályt különítettek el: a hangsúly-időzítésű, a szótag-időzítésű és a moraidőzítésű nyelveket (Pike 1945; Abercombie 1967). Feltételezték, hogy az különbözteti meg a nyelveket egymástól, hogy más a beszédritmus alapegysége. A hangsúly-időzítésű nyelvek esetében a hangsúlytól hangsúlyig terjedő szakaszokról gondolták, hogy közel azonos hosszúságúak. A szótag-időzítésű nyelvekben a szótagok időtartamáról, a moraidőzítésű nyelvekben pedig a morák (azaz a szótag súlyát megadó fonológiai egységek) időtartamáról tartották úgy, hogy nagyjából hasonló időtartamban realizálódnak (Pike 1945; Abercombie 1967; Ladefoged 1975). Az akusztikai vizsgálatok nem támasztották alá ezeket a hipotéziseket (Roach 1982), de a beszédritmus-mérőszámok és a felállított nyelvi osztályok között összefüggéseket találtak. A magánhangzós és a mássalhangzós szakaszok szórása, valamint a PVI mérőszámok is tendenciaszerűen magasabb értékeket mutattak a tipikusan hangsúly-időzítésűekhez sorolt nyelveknél, mint a tipikusan szótag-időzítésűekhez sorolt nyelveknél. A mérőszámok által kapott eredmények azonban nem vezetnek egyértelmű csoportosításhoz, helyette a nyelvek beszédritmusát a főbb kategóriák közti átmenetekként határozhatjuk meg (Ramus et al. 1999; Grabe–Low 2002).

A beszédritmus-mérőszámok megbízhatóságával kapcsolatban azonban több probléma is felmerült. Az egyik nehézség, hogy a beszédritmus-mérőszámok egy része összefüggést mutatott az artikulációs tempóval. Ennek megoldására az artikulációs tempótól függetlennek tekinthető mérőszámokat (VarcoV, VarcoC) hoztak létre, amelyek azonban hasonló eredményeket hoztak a korábbiakhoz képest (l. Dellwo 2009). További problémát jelent, hogy a beszélők közti nagyfokú variabilitás elmoshatja a különbségeket az alapvetően eltérő beszédritmusú nyelvek között. Szintén nehézséget jelent, hogy a mérőszámok érzékenyek a vizsgált anyag felépítésére, beszédmódjára is (Arvaniti 2012). Az eddigi elemzések alapján tehát nem teljesen egyértelmű, hogy a nyelvek közti beszédritmusbeli különbségek mennyire megbízhatóan határozhatók meg a mérőszámok alapján (Arvaniti 2012). A megközelítés módszertana és alapegységei azonban hasznos eszközök lehetnek nyelveken belüli időbeli változások vizsgálatára.

A beszédritmus-mérőszámokon kívül más megközelítések is születtek a folyamatos beszédben lévő hangok időtartamának megragadására. A beszéd gépi előállításához az időszerkezet modellezése elengedhetetlen feladat. Az egyik magyar beszéd szintézis-modell kiindulópontja az ún. specifikus időtartam. A specifikus időtartam olyan alapidőtartamot jelent, amely a beszédhang artikulációs konfigurációjából és a közvetlen hangkörnyezetéből fakad (Olaszy

2010). A specifikus időtartamot ebben a modellben percepciók ítéletek alapján határozzák meg, majd az így kapott eredményeket szabályok révén módosítják a végleges időtartam elérése céljából. A specifikus időtartamokat először a szó szintjén változtatják meg. A szó hosszától függően 6 szótagig rövidítik a hangokat. A mondat szintjén veszik figyelembe a hangsúlyozást. A hangsúlyos helyzetben a hangzókat nyújtják, hangsúlytalan szótagokban rövidítik, ami egyfajta gyorsítást-lassítást hoz létre az időszerkezetben (l. Olaszy 2006). További szabály, hogy minél hosszabb egy mondat, annál jobban rövidítik a benne lévő hangok időtartamát. A modell a szó helyzetét is figyelembe veszi a mondaton belül, a szünet előtti szót, különösen a szónak a végét lassítják. A szünetet határoló elemnek tekintik az időszerkezet tekintetében, a szünet után ugyanis mindig új időszerkezeti frázis indul. Az e modell segítségével előállított beszéd időtartamadatai hasonló mintázatot mutattak a természetes felolvasásból származó időtartamadatokkal, ami a meghatározott szabályok, szabályszerűségek meglétére utal (Olaszy 2006, 2010).

Más módszerrel készült szövegfelolvasó programok további információkkal szolgálhatnak a beszéd időszerkezetéről. Az egyik magyar nyelvű, rejtett Markov-modellen alapuló gépi szövegfelolvasó szerint a hangzók képzési konfigurációja alapvetően meghatározza az időtartamokat. A modell felépítéséhez további olyan tényezőket vesznek figyelembe, mint például a szótag hangsúlyos vagy hangsúlytalan volta, a szótagok száma az előző vagy a következő hangsúlyos szótagtól, a fonológiai felépítettség, az aktuális szótag pozíciója a szóban, a szó pozíciója a mondatban, a szavak száma a mondatban stb. (Tóth 2013).

Az eddigi modellek, megközelítések a beszédhangok időtartamát befolyásoló összes paraméter együttes hatásának megragadására törekedtek. Az időtartamokkal foglalkozó akusztikai kutatások egy jelentős része viszont a paraméterek hatását külön-külön vizsgálja. A közlésben megjelenő különböző pozíciók és a hangzók időtartama között például több összefüggést is kimértek. A hangsorban elfoglalt hely hatással lehet a magánhangzók időtartamára. A hangsor végi nyúlást több nyelvben (pl. angolban, franciában, oroszban, csehben) mérték, de a lassítás mértéke és hatóköre a különböző nyelvekben eltérőnek bizonyult (Hardcastle et al. 2010). Magyar nyelvű felolvasásokban a magánhangzók hosszabbnak bizonyultak abszolút hangsor végi helyzetben, mint közlés közben (Magdics 1966; Kassai 1979). A felolvasásokban nemcsak abszolút hangsorvégen, hanem CVC felépítésű utolsó szótagban is nyúlnak a magánhangzók, akár rövid, akár hosszú magánhangzókról van szó (White–Mády 2008). A hosszú magánhangzók az utolsó előtti szótagban is megnyúlnak (White–Mády 2008), Kovács (2002) vizsgálatai alapján a rövid magánhangzók is kimutatható az időtartambeli növekedés az utolsó előtti szótagban. A magánhangzók mellett a mássalhangzókat is hosszabbnak találták a felolvasások hangsor végi pozíciójában (Magdics 1966; Kassai 1979). A hangsor végi pozíció mellett a közlés kezdete is kitüntetett lehet, a magyar-

ra végzett kutatások eredményei azonban ellentmondásosak. Magdics (1966) szerint mind a magánhangzók, mind a mássalhangzók rövidebbek felolvasott mondatok kezdetén, mint hangsor belsejében. Kassai (1979) szerint viszont a hangsor eleji hangzók hosszabbak a közlés belseji hangzóknál. Az eltérések fakadhatnak az eltérő anyagból és szemléletből. Kassai (1979) a hangsor eleji pozíciót úgy határozta meg, hogy a felolvasások mondatkezdő vagy szünet utáni hangzója, míg más kutatásokban csak mondatkezdő pozíciókat vizsgáltak (Magdics 1966; White–Mády 2008). A hangsor eleji pozíció időtartamra gyakorolt hatása tehát nem egyértelmű, de a hangsor végi lassítás meglétét a gépi felolvasó rendszerek modelljei mellett a felolvasásokon végzett akusztikai mérések is megerősítik.

A közlésbeli pozíciókon kívül más tényezők is kiválthatnak lassuló vagy gyorsuló szakaszokat a beszédben. A szóhosszúság növekedésének hatására a hangsúlyos helyzetű magánhangzók időtartama megrövidülhet (pl. angolra: Lehiste 1972; Turk–Shattuck–Hufnagel 2000). A magyar magánhangzókra végzett kísérletek azt mutatták, hogy a magánhangzók időtartama rövidebb hosszabb szavakban, mint rövidebb szavakban (Meyer–Gombocz 1909; Tarnóczy 1974; Kassai 1979). Ennek a jelenségnek a meglétét azonban nem minden kutatás támasztotta alá (vö. White–Mády 2008). Lassulásokat és gyorsulásokat okozó további tényező lehet például a hangsúlyos pozíció, de a magyar nyelven végzett akusztikai mérések nem mutatnak egyértelmű összefüggést a perceptuálisan észlelhető hangsúlyos helyzet és az időtartamok között (pl. Markó 2012). A hangsúlyos helyzet érzetét ugyanis más paraméterek (pl. intenzitás, alaphangmagasság) változása is keltheti, amely nem feltétlen jár együtt időtartambeli eltéréssel.

A lassuló és gyorsuló szakaszok éppen aktuális kiváltó okait ugyan nem feltétlen ismerjük, de az időtartambeli folyamatos váltakozások mérésére már több kísérlet is született. A lassulások és a gyorsulások előfordulásának aránya nem szimmetrikus. A magyar nyelvben az intonációs frázisokon belül a szavak tendenciaszerűen lassulnak az egységek vége felé (Váradi–Beke 2013). Az egymást követő magánhangzós és mássalhangzós szakaszok időtartamának vizsgálatai is azt mutatták, hogy lassuló tendencia figyelhető meg a tagmondatnyi egységeken belül (Kohári 2013). A lassulások és gyorsulások azonban nem alkotnak hosszabb tömböket. Ezt támasztja alá, hogy a szótagoknak egyfajta időtartambeli váltakozása mutatható ki mind normál ejtésű, mind reppelt szövegek esetében (Gósy 2000). A szótagokon kívül az egymást követő beszédhangok vagy a magánhangzós és a mássalhangzós szakaszok sem alkotnak folyamatosan növekvő vagy folyamatosan csökkenő szakaszokat (Kohári 2013). A lassulások és gyorsulások határai tehát nem határozhatók meg egyértelműen a beszédhangok vagy magánhangzós és mássalhangzós szakaszok vagy szótagok időtartamai alapján, önmagukban, más paraméterek nélkül nem szolgálnak információval a temporális változások pontos helyéről. Az artikulációs tempót vizsgáló kutatások ugyanakkor azt mutatják,

hogy az időbeli mintázatok eltérőek lehetnek attól függően, hogy mekkora alapegységet vizsgálunk (l. de Looze 2010). Az intonációs frázisokban például az artikulációs tempó nagyobb variabilitása jellemző, mint a hosszabb beszédszakaszokban (szünettől szünetig tartó egységekben) (Olaszy 2006). Ha az artikulációs tempó jellemzői változnak a vizsgált szakasz nagyságától függően, akkor felmerül a kérdés, hogy az intonációs frázisokra és a tagmondatnyi egységekre jellemző lassuló trend hosszabb szakaszok esetében hogyan realizálódik. Amennyiben egy sok beszédszakaszból álló szöveget próbálnánk vizsgálni, akkor a benne lévő beszédszakaszok általános lassuló trendjéből nem következne, hogy a teljes szöveg is lassuló trendet mutat.

A jelen kutatás célja, hogy kísérletet tegyen a lassulások és gyorsulások jellemzőinek megismerésére beszédszakaszoknál nagyobb egységekben. Kutatásunkban különböző hosszúságú egységek lassulásainak és gyorsulásainak mértékét és arányait próbáljuk összehasonlítani egy leíró statisztikai módszer segítségével akusztikai mérések alapján. Feltételeztük, hogy a lassulások és gyorsulások eltérő arányban, módon realizálódnak a beszédszakaszok sokaságából összeálló szövegekben, mint az egyes beszédszakaszok szintjén.

Anyag, módszer, kísérleti személyek

Kutatásunkhoz 10 magyar anyanyelvű beszélő (5 nő és 5 férfi) spontán beszédét és szövegfelolvasását vizsgáltuk a BEA adatbázisból (Gósy 2008). Életkoruk 20 és 60 év között mozgott (a férfiak átlagéletkora: 41,6 év, a nőké: 38,0 év). Az adatközlők ép hallók és ép beszédűek voltak. Interjúhelyzetben rögzített spontán beszédükből beszélőnként legalább 2 percnyi részletet választottunk ki, melyben az interjúztató nem szólalt meg. A beszélők saját munkájukról vagy éppen folyó tanulmányaikról monologikusan beszéltek a felvételeken. Továbbá megvizsgáltuk mindegyik beszélő szövegfelolvasását is. Ugyanazon szöveg felolvasása a különböző beszélők esetében eltérő időtartamban valósult meg, a legrövidebb szövegfelolvasás 107,7 másodperc, a leghosszabb 159,7 másodperc volt. A hanganyagot hangszinten annotáltuk a magyar nyelvre is adaptált MAUS elnevezésű automatikusan szegmentáló szoftverrel (Schiel 1999). Jelöltük a beszédhangok minőségét, határaikat, valamint a szüneteket. Az így kapott hanghatárokat manuálisan ellenőriztük a Praat 5.3 szoftverben (Boersma–Weenink 2013). A magánhangzók időtartamát a formánsszerkezet kezdetéhez és végéhez igazítottuk az oscillogram, a spektrogram és auditív információk segítségével, követve a nemzetközi szakirodalomban szokásos hangelhatárolási kritériumokat (Grabe–Low 2002). Azokban az esetekben, ahol a hanghatárok nem voltak egyértelműek az oscillogram és a spektrogram alapján (pl. magánhangzó nazális mássalhangzó vagy approximáns környezetében), a hanghatárt a formánsátmenet felénél helyeztük el. A közlést kezdő és a szünet utáni zöngétlen felpattanó zárhangok zárszakaszának idejét úgy határoztuk meg, hogy ugyanazon közlés következő, ugyanolyan képzésű hangzó zárszakaszának idejét hozzáadtuk a

zárhang zörejes részéhez. A hangszintű címkesor alapján kiszámítottuk a magánhangzós és a mássalhangzós szakaszok időtartamát. A hiátustöltő funkciójú *j*-realizációkat a magánhangzós szakasz részének tekintettük. Ily módon a torlódott magánhangzók időtartama, hiátustöltéssel együttesen adta az adott magánhangzós szakasz időtartamát. A szakaszok átívelték a szóhatárokon és szüneteken, de a szünetek nem képezték részét a szakaszok időtartamának (hasonlóan: Grabe–Low 2002). A magánhangzós szakaszok és a mássalhangzós szakaszok idejét egy C++ nyelven írt program segítségével számoltuk ki.

A temporális változások vizsgálatához az ún. lépésszatisztika módszert alkalmaztuk (Gyüre et al. 2007). A korábban más tudományágban használt módszer célja, hogy hosszú idősorok szimmetriáját vagy aszimmetriáját felfedje. A módszer a beszéd folyamatban lévő lassulások és gyorsulások arányának feltárására a következőképpen alkalmazható. Az egymást követő egységek (beszédhangok, magánhangzós szakaszok és mássalhangzós szakaszok) időtartamának különbségét rendre kiszámoljuk. Minden egyes különbség egy-egy ún. lépést jelent az egymás mellett lévő egységek között. Amennyiben a követő egység időtartama hosszabb, mint a megelőző egységé, akkor a különbséget növekvő lépésnek tekintjük, mivel ez egyfajta lassulást jelez. Amennyiben a követő egység időtartama rövidebb, mint a megelőző egység időtartama, akkor ez felfogható gyorsulásnak, ezért csökkenő lépésként kezeljük a továbbiakban. Az 1. táblázat azt szemlélteti, hogy az egymást követő egységek időtartamának különbségeit hogyan minősítjük növekvő vagy csökkenő lépéseknek. Ebben az esetben az alapegységek az egymást szükségszerűen felváltva követő mássalhangzós szakaszok és magánhangzós szakaszok (CV-szakaszok).

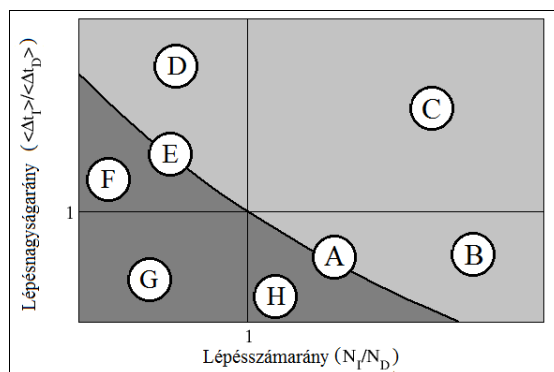
1. táblázat: Példa a csökkenő (N_D) és növekvő (N_I) lépések meghatározására (I_v a magánhangzós szakaszt, I_c a mássalhangzós szakaszt jelöli.)

Egymást követő alapegységek	I_v	I_c	I_v	I_c
Mért időtartam (ms)	50	40	60	45
A különbségek	-10	+20	-15	
Lépés típusa	N_D	N_I	N_D	

A példában az első egység egy magánhangzós szakasz, melynek időtartama (50 ms) hosszabb a követő mássalhangzós szakasz időtartamánál (40 ms), tehát az első egység időtartamához képest a következő egység rövidebb, így az első lépés típusa csökkenő lesz. A második egység (mássalhangzós szakasz) időtartamához képest a harmadik egység (magánhangzós szakasz) időtartama több, így az egy növekvő lépés lesz. A kategorizálás után megszámoljuk külön a növekvő (N_I) lépéseket és külön a csökkenő lépések (N_D) darabszámát. A növekvő lépések darabszámát elosztjuk a csökkenő lépések darabszámával, tehát a kettő aránya (N_I/N_D) adja az ún. lépésszamarányt. A ms-

ban mért időtartamokat négy tizedesjegyre vettük figyelembe, és a teljesen megegyező időtartamokat (2 db) a növekvő időtartamokhoz számítottuk. Mivel a különbségek nagyságukat tekintve rendkívül kicsik is lehetnek, ezért önmagában a lépésszámarányok nem szolgáltatnának információval a trendek relevanciájáról. Ezért nemcsak a lépések számát, hanem a nagyságát (az időtartam-különbségek abszolút értékét) is figyelembe vesszük. Ehhez kiszámoljuk külön a növekvő lépések átlagos nagyságát ($\langle \Delta t_i \rangle$), valamint a csökkenő lépések átlagos nagyságát ($\langle \Delta t_D \rangle$). A növekvő lépések átlagos nagyságát elosztjuk a csökkenő lépések átlagos nagyságával, tehát a kettő aránya határozza meg az átlagos lépésmagyságarányt ($\langle \Delta t_i \rangle / \langle \Delta t_D \rangle$). Az átlagolt szakasz hosszának a teljes szöveget tekintettük, illetve összevetésképpen megvizsgáltuk külön a beszédszakaszokra kapott értékeket is.

Az átlagos lépésmagyságarány és a lépésszámarány olyan paramétersíkot jelöl ki, amelyben jellemezhetővé válik a gyorsulások és lassulások számának és mértékének egymáshoz viszonyított aránya (1. ábra). Ha mindkét arányszám pontosan 1, akkor az idősor tökéletesen szimmetrikus. Azon pontok halmaza a paramétersíkon, melyek a lépésszám- és lépésmagyságarányok olyan viszonyával jellemezhetők, hogy csökkenő vagy növekvő trend nélküli idősort írnak le, azok az $\langle \Delta t_i \rangle / \langle \Delta t_D \rangle = N_D / N_I$ görbén (azaz egy $y = 1/x$ jellegű hiperbolán) helyezkednek el. A paraméterter ezen tartományát stacionárius görbének nevezzük. A stacionárius görbe tehát olyan pontok összességéként definiálható, amelyek esetében igaz az, hogy a gyorsulások és a lassulások kiegyenlítik egymást, mértékük és gyakoriságuk fordítottan arányos. Az 1. ábrán a stacionárius görbét vastag fekete vonallal jelöltük. A görbe fölött elhelyezkedő pontok lassulást mutatnak, míg a görbe alatt elhelyezkedő pontokra a gyorsulás jellemző.



1. ábra

Szemléltető ábra a lépésstatistikai módszer lényegéről lassuló trendet mutató adatok tartománya (világosszürke), gyorsuló trendet mutató adatok tartománya (sötétszürke), stacionárius görbe (fekete vonal)

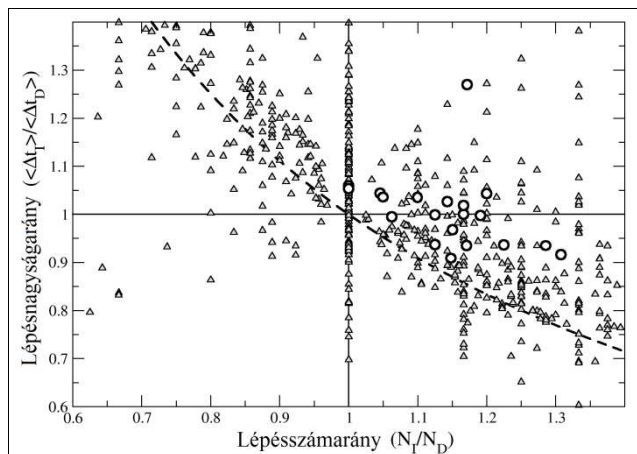
Az alábbiakban végigvesszük a különböző lehetséges eseteket. Az *A* esetben a több növekvő lépés kompenzálja a kevesebbszer előforduló, ám átlagosan nagyobb csökkenő lépéseket. A *B* esetben a sokkal gyakoribb növekvő lépéseket a csökkenő lépések nagysága már nem tudja ellensúlyozni, így az ezen a területen szóródó adatok lassuló trendet mutatnak. Amennyiben mindkét mérőszám nagyobb 1-nél (*C* eset), akkor a növekvő lépések gyakorisága és nagysága is nagyobb a csökkenő lépéseknél, tehát ez esetben egyértelmű lassuló trendről beszélhetünk. Előfordulhat az is, hogy a lépésszámarány kisebb 1-nél, de a lépésnagyságarány nagyobb, mint 1 (*D*, *E*, *F* eset). *D* esetben az átlagosan nagyobb növekvő lépéseket a gyakoribb csökkenő lépések nem tudják kompenzálni, így itt lassulás következik be. Fordított a helyzet az *F* esetben, ahol a gyorsulás a jellemző, ugyanis a gyakoribb csökkenő lépéseket nem tudják ellensúlyozni az átlagosan nagyobb növekvő lépések. *E* esetben a gyakoribb, ámde átlagosan kisebb csökkenő lépések kompenzálják a ritkább, de átlagosan nagyobb növekvő lépések, tehát hasonlóan az *A* esethez itt sincs sem lassuló, sem gyorsuló trend. A *G* esetben egyértelmű gyorsulás figyelhető meg, mert a csökkenő lépések átlagos mértéke és gyakorisága is nagyobb a növekvő lépéseknél. *H* esetben az átlagosan nagyobb csökkenő lépéseket a gyakoribb növekvő lépéseket nem tudják kompenzálni, ezért itt is gyorsulásról beszélhetünk (l. Kohári 2013).

Az eddig leírt módszer alapvetően három nagy csoportot hoz létre, a lassuló trendhez tartozók, gyorsuló trendhez sorolhatóak és a trend nélküliek csoportját. A trend nélküliek csoportja nem területtel, hanem egyetlen görbével jellemezhető, ezért a gyakorlatban kevés a valószínűsége, hogy a mért adatok ebbe a csoportba kerüljenek. Ennek kiküszöbölésére szélesebb sávot határoztunk meg, és az e sávba eső adatpontokat tekintettük trend nélkülinek. Először a mért időtartam adatok sorrendjét megváltoztattuk, véletlenszerűen összekevertük egy nyílt forráskódú Time Series Analysis (TiSeAn) 3.0.1 nevű szoftvercsomaggal (Schreiber–Schmitz 2000). A keverés során egy véletlenszám-generátor az időtartam adatokhoz új sorszámokat rendelt. Ezzel a módszerrel 10 különböző kevert idősort hoztunk létre, amelyekre szintén kiszámoltuk a lépésszatisztikát. Meghatároztuk az így kapott eredmények stacionárius görbéhez viszonyított eltéréseinek maximumát (Δ_{\max}), majd a stacionárius görbét mindkét irányban eltoltuk ezzel az értékkel. Vizsgálataink szempontjából akkor tekintettünk valamely adatpontot a stacionáriusgörbén fekvőnek, ha annak koordinátái a görbét leíró $\langle \Delta t_i \rangle / \langle \Delta t_D \rangle$ értékek $\pm \Delta_{\max}$ szélességű környezetébe estek.

Eredmények

A 10 beszélőtől összesen 1079 beszédszakaszt (szünettől szünetig tartó szöveges egységet) elemeztünk. A beszédszakaszok igen nagy szórást mutatnak a lépésszatisztika alapján. A lépésszatisztika alapegységeként először a magánhangzós és mássalhangzós szakaszokat (CV-szakaszok) együttesen

vizsgáltuk. Vannak olyan beszédszakaszok, amelyeket inkább a lassulás, másokat inkább a gyorsulás jellemez (2. ábra). Az összes beszédszakaszt figyelembe véve azonban kirajzolódik egy általános lassuló trend. Az egyes beszélők beszédszakaszainak külön a spontán beszédre és külön a szövegfeldolvasásra vett mediánértékei is ezt támasztják alá, ugyanis a mediánokat jelölő adatpontok mindegyike az ún. stacionárius görbe fölött helyezkedik el. A növekvő lépések száma több vagy ugyanannyi, mint a csökkenő lépések száma, tehát a lépésszámarány nagyobb vagy egyenlő 1-gyel. Az átlagos lépésmagyságarány 1 körül szóródik (0,908–1,270). A csökkenő lépések átlagos nagysága ugyan olykor nagyobb, mint a növekvő lépéseké, a növekvő lépések gyakoribb előfordulása elbillenti a mérleget a lassulás felé. A beszédszakaszok többsége tehát lassuló tendenciát mutat, egy kisebb részüket azonban gyorsuló trend jellemzi.

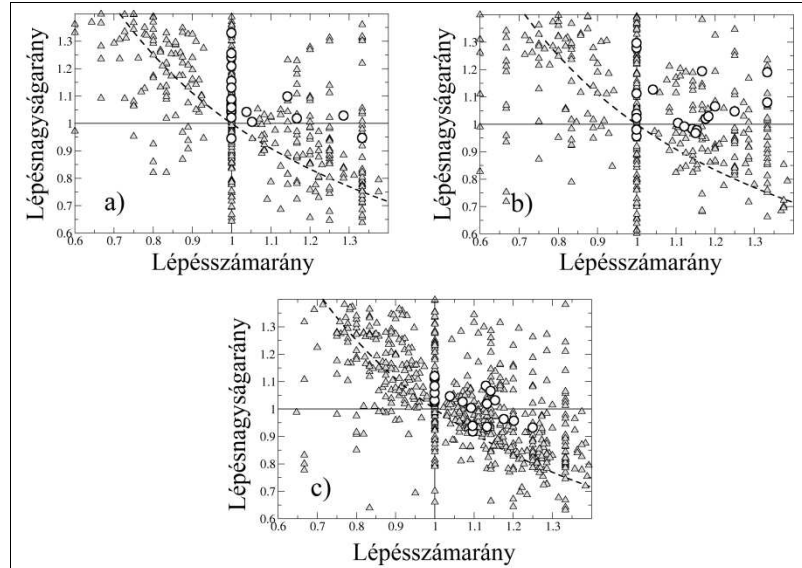


2. ábra

A CV-szakaszok beszédszakaszaira kapott lépésszatisztikai eredmények (háromszög) és a beszélők különböző beszédmódjaira vett mediánok (kör)

A CV-szakasz egy önkényesen kiválasztott lehetséges alapegysége a lépésszatisztikának, ezért megvizsgáltuk más, a szakirodalomban megtalálható alapegységekre (beszédhangok, V-szakaszok, C-szakaszok) is a gyorsuló és lassuló trendeket. Mindegyik alapegység esetén hasonló eredményeket mutat a lépésszatisztika a beszédszakaszokra. Ha a beszédhangok közti különbségeket vesszük alapul, akkor a beszédszakaszok nagy területen szóródnak (3. c) ábra). A beszélők két beszédmódjára vett mediánok szintén lassuló tendenciát mutatnak, mivel a stacionárius görbe fölött helyezkednek el (lépésszámarány: 1,000–1,250, átlagos lépésmagyságarány: 0,918–1,119). A magánhangzós szakaszok időtartamára számolt lépésszatisztika (3. b) ábra), illetve külön a

mássalhangzós szakaszok időtartamára számolt lépésstatistika (3. a) ábra) is nagy szórást mutat beszédszakaszokra számolva. A jellemzően lassuló és az inkább gyorsuló beszédszakaszok közül az előbbiek vannak túlsúlyban mindkét esetben. Ezt jelzik a beszélőkre számolt mediánok, amelyek mind a mássalhangzós szakaszok esetében, mind a beszédszakaszok esetében tendenciaszerűen a stacionárius görbe fölött szóródnak. A beszédszakaszokra tehát alapegységtől függetlenül a lassulás jellemző, annak ellenére, hogy előfordulnak jellegzetesen gyorsuló beszédszakaszok is.

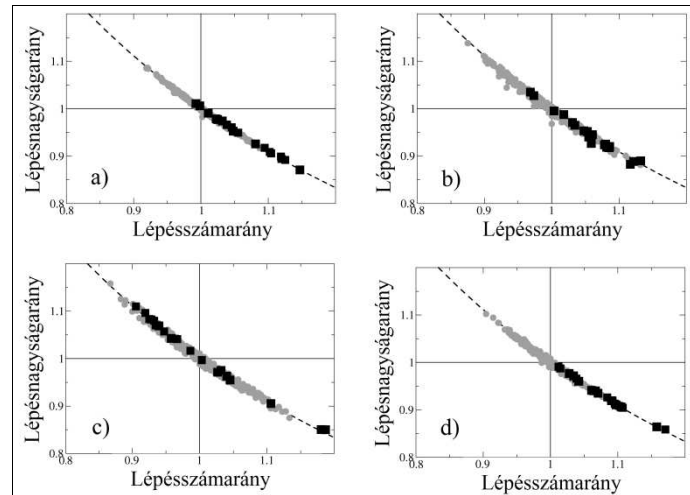


3. ábra

Az a) C-szakaszok, b) V-szakaszok és c) beszédhangok beszédszakaszaira kapott lépésstatistikai eredmények (háromszög), valamint a beszélők különböző beszédmódjaira vett mediánok (kör)

A beszédszakaszok vizsgálata után megváltoztattuk az elemzés időablakának méretét, és kiszámoltuk a lépésstatistikát a beszélők teljes spontán beszédére és teljes szövegfelolvasására is. Az eredmények jobb láthatósága érdekében eltér a beszédszakaszok és a hosszabb szövegek lépésstatistikáját szemléltető ábra léptéke, de a pontok szóródásának különbsége szembevető. A lassulások és a gyorsulások mértéke, illetve gyakorisága kiegyenlített a hosszabb szövegek esetén, bármelyik alapegységet (beszédhangok, C-szakaszok, V-szakaszok, CV-szakaszok) vesszük alapul a lépésstatistikához (4. ábra). Az egyes beszélők hosszabb szövegeire számolt lépésstatistikai eredmények a stacionárius görbén szóródnak. Az adatpontok hibahatáron belül, a

vonalon találhatók a Monte Carlo-teszt alapján (Schreiber–Schmitz 2000). A CV-szakaszok esetében minden mért adatpontra számolt lépésstatistikai eredmény a kevert adatsorra számolt eredmények alapján meghúzott sávon belül helyezkedik el ($\Delta_{\max} = 0,014$). Az eredmények mindegyike a stacionárius görbén elhelyezkedőnek tekinthető a statisztikai próba alapján akár a C-szakaszokat ($\Delta_{\max} = 0,013$), akár a V-szakaszokat ($\Delta_{\max} = 0,032$), akár a beszédhangokat ($\Delta_{\max} = 0,016$) vesszük alapegységnek.

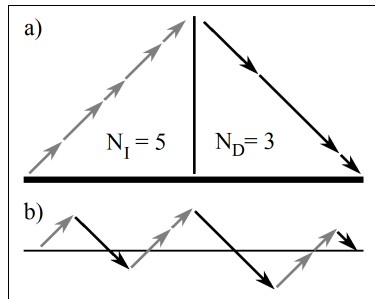


4. ábra

A hosszú szövegekre (fekete négyzet) és a kevert adatsorokra (szürke kör) kapott lépésstatistikai eredmények *a)* CV-szakaszokra, *b)* C-szakaszokra, *c)* V-szakaszokra és *d)* beszédhangokra

A hosszabb szövegek tehát kiegyenlítettek a lassulás és gyorsulás szempontjából, de nem teljesen szimmetrikusak. A kevert adatsorokra számolt lépésstatistikai eredményekből is látszik, hogy egy struktúra nélküli idősor esetében a lépésszámarány egyaránt lehet 1-nél nagyobb, és kisebb is (0,867–1,132). A mért adatsorok viszont nem szórnak a teljes tartományban. A CV-szakaszok időtartamára kapott lépésszámarányok minden esetben nagyobbak 1-nél (1,012–1,171). A beszédhangok esetében a lépésnagyságok két adat kivételével nagyobbak 1-nél (0,993–1,147). A magánhangzós szakaszok esetében hasonlóan alakul a lépésnagyság (0,968–1,113), mindössze két adat kisebb 1-nél. A mássalhangzós szakaszok esetében nem figyelhető meg hasonló tendencia a lépésnagyság tekintetében (0,905–1,187). A CV-szakaszok, a beszédhangok és a V-szakaszok esetében tehát olyan mintázat látható, amelyben a növekvő lépések száma valamelyest nagyobb a csökkenő lépéseknél,

de a csökkenő lépések átlagos nagysága nagyobb a növekvő lépéseknél. A növekvő lépések gyakoribb előfordulását a csökkenő lépések nagyobb mértéke ellensúlyozza, ahogyan azt az 5. a) szemléltető ábra mutatja. A beszéd folyamatában az egymást követő beszédhangok vagy magánhangzós és mássalhangzós szakaszok időtartama azonban nem alkot folyamatosan növekvő vagy folyamatosan csökkenő tömböket. A növekvő és csökkenő lépések váltogatják egymást (5. b) ábra).



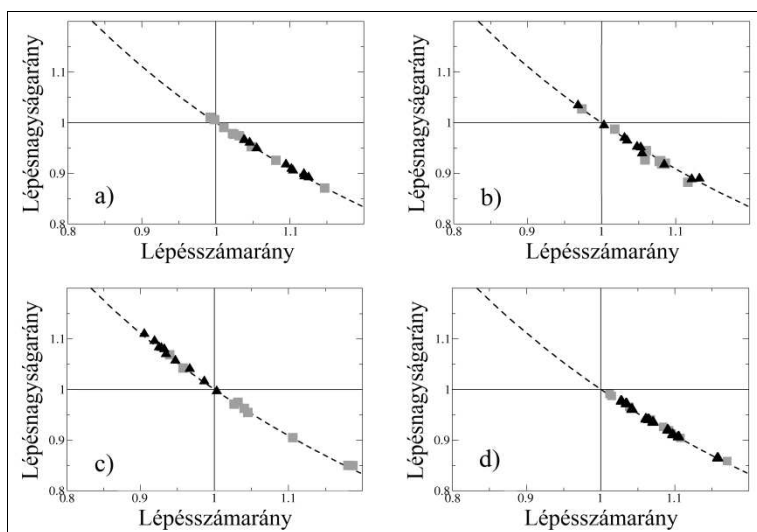
5. ábra

Az a) a növekvő lépésekre (szürke felfelé mutató nyíl) és a csökkenő lépésekre (fekete lefelé mutató nyíl) talált mintázat egy lehetséges példája és b) ugyanannak a mintázatnak időbeli sorrendben megvalósuló példája

A különböző beszéd módok hatással lehetnek a hangzók időtartamaira, illetve az időtartamok variabilitására, ezért megvizsgáltuk külön a szövegfelolvasásra és külön a spontán beszédre kapott lépésszisztemikai eredményeket. Az elemzés időablaka ezúttal is a teljes felolvasás, illetve a teljes spontán beszéd-részlet volt. A beszédhangokra, V-szakaszokra, CV-szakaszokra számolt lépésszámok erős átfedést mutatnak a beszéd módok között (6. ábra). A C-szakaszok esetében viszont a szövegfelolvasásra kapott lépésszám arányok egy kivétellel 1 alatt szóródnak (0,905–1,003). Ez azt jelenti, hogy a szövegfelolvasásokban a C-szakaszok közötti csökkenő lépések gyakoribbak, ám valamivel átlagosan kisebbek, mint a növekvő lépések. A spontán beszédre kapott arányok azonban nem mutatnak ilyen tendenciát (0,940–1,187). Amennyiben tehát a C-szakaszokat tekintjük alapegységnek, a kétféle beszéd mód-ban eltérő temporális mintázat jelenik meg.

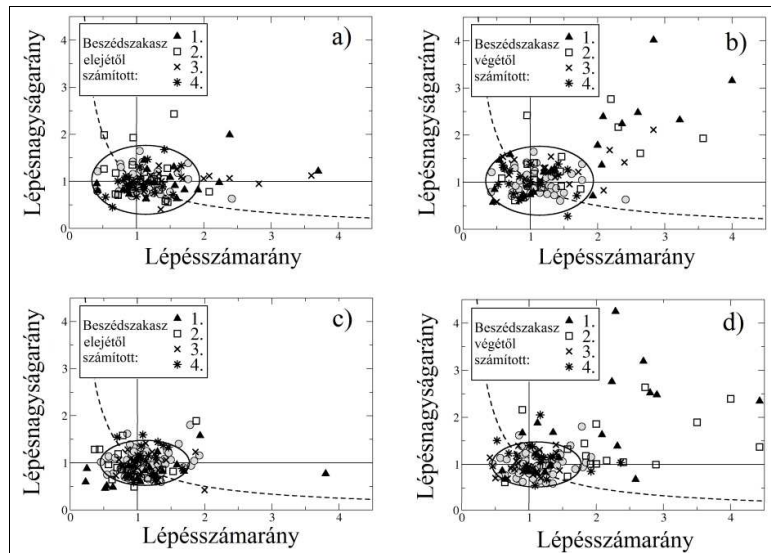
A hosszabb szövegekben kiegyenlítettebb a lassulások és gyorsulások mértéke és gyakorisága, mint a beszéd szakaszok esetében, de nem teljesen szimmetrikus. A lassulás és gyorsulás különböző mintázatainak magyarázataként felmerülhet, hogy a közlések különböző pozíciói (hangsor eleji, hangsor végi helyzet) szerepet játszanak az aszimmetrikus elrendeződés kialakulásában. Spontán beszédben a tagmondatok határai nehezen határozhatók meg, így a szünetek előtti és utáni pozíciókat vizsgáltuk meg. Mivel a spontán beszéd-

ben nem határozhatók meg egyértelműen gondolati egységek vagy a felolvasásokhoz hasonló mondatszerű egységek (Markó 2010; Váradí 2013), ezért a szünetek különböző funkcióinak megállapításától és elkülönítésétől eltekintettünk. A szünetek előtti és utáni pozíciók vizsgálatakor a magánhangzós és a mássalhangzós szakaszokat csak a szünettől szünetig tartó szakaszokon belül értelmeztük. Az összehasonlíthatóság miatt csak azoknak a lépéseknek az előfordulásait vettük figyelembe, amelyek után legalább kilenc lépés következett. Ennek oka, hogy a hangsor végi nyúlás akár az utolsó előtti magánhangzón is megfigyelhető, melynek CV-szakaszban való lefedése négy egységet jelent, ugyanennyi egységet számoltunk a beszédszakasz kezdetén jelentkező jelenségek maximális hatókörének. A beszédszakasz belseji helyzetre szabtuk egy minimumot (3 lépés), amelynek összességét tekintettük a másik két pozíció viszonyítási alapjának (7. ábra szürke kör). Ezekre a viszonyítási pontokra kiszámítottuk mind a lépésszámarány, mind az átlagos lépésmagasság átlagát és szórását. A 7. ábrán látható ellipszisek középpontját az átlagok adják meg, az ellipszis vízszintes és függőleges tengelyének a hossza a szórások négyszerese minden esetben. Az így kapott ellipszis a viszonyítási pontok legalább 90%-át tartalmazzák.



6. ábra

A spontán beszédre (szürke négyzet) és a felolvasásra (fekete háromszög) számolt lépésstatistikai eredmények a) CV-szakaszokra, b) C-szakaszokra, c) V-szakaszokra és d) beszédhangokra



7. ábra

A beszédszakasz a) eleji és b) végi pozíciók CV-szakaszokra, valamint a beszédszakasz c) eleji és d) végi pozíciók beszédhangokra számolt lépésstatisztikája

A beszédhangok esetében ugyanezt az eljárást alkalmaztuk. A különböző pozícióhoz köthető lépéseket aszerint csoportosítottuk, hogy melyik beszélő mely beszédmódjához tartozott, majd a csoportra külön-külön kiszámoltuk a lépésstatisztikát. A beszédszakasz belseji helyzetek CV-szakaszokra kapott lépésnagysága 0,604–1,642 között, lépésszáma 0,417–2,147 között szóródott. A beszédhangokra kapott lépésszámárány 0,458–1,923 között, a lépésnagyságárány pedig 0,803–1,593 között mozgott.

A beszédszakaszok elején lévő lépések arányszámai azon a területen szóródnak, ahol a beszédszakaszok belsejében lévő lépések helyezkednek el, akár a CV-szakaszokat, akár a beszédhangokat vesszük alapul. A szünet előtti pozíciókban lévő lépések arányszámai nagyobb szórást mutatnak, mint a beszédszakasz belseji pozícióban lévő lépéseké. Mind a CV-szakaszok, mind a beszédhangok esetében előfordulnak olyan realizációk, amelyek egyértelmű nyújtást mutatnak (7. ábra). Mivel a csak C- és csak V-szakaszok ugyanilyen hatókörű vizsgálata a különböző pozíciók elemzését csak kevés elemszámmal tette volna lehetővé, ezért a lépésstatisztikai módszerekkel való tanulmányozásuktól eltekintettünk.

Az egymást követő időtartamok különbségeinek vizsgálata egy sajátos pozíció vizsgálatára is lehetőséget ad. A szünet előtti utolsó egység, illetve a szünet

utáni első egység időtartamának különbsége ugyanis vizsgálhatóvá válik a szüneteken átívelő lépések segítségével. A többi pozícióval való összehasonlíthatóság megtartásának céljából csak azokat az előfordulásokat vettük figyelembe, amely esetekben a fentebb definiált módon egyértelműen szünet előtti utolsó pozíciójú és szünet utáni első pozíciójú egység időtartam-különbségeit tudtuk kiszámolni. A szüneteken átívelő lépéseknek a beszélők hosszú szövegeire külön számolt lépésszisztektikája egyértelmű gyorsulást mutat. A lépésszámarány minden beszélő esetében messze elmarad az 1-től (CV-szakaszokra: 0,053–0,562, beszédhangokra: 0,075–0,733). A lépésmagyságarány a lépésszámarányhoz hasonlóan minden beszélő adataiban alacsonyabb 1-nél (CV-szakaszokra: 0,137–0,829, beszédhangokra: 0,016–0,963). A szünetek előtti utolsó beszédhangot vagy szakaszt tehát gyakran követi a szünet után egy jóval rövidebb időtartamú egység.

Következtetések

Tanulmányunkban a beszéd időbeli változásait az egymást követő beszédhangok és szakaszok időtartamainak különbsége felől közelítettük meg egy, a beszédhangokra adaptált leíró statisztikai módszer segítségével. Az eredmények alapján elmondható, hogy a vizsgált beszédész méretének változtatásával más időbeli mintázatok jelentek meg. A beszédszakaszok rendkívül változatosan realizálódtak, de jellemzően lassuló trendet mutattak. A beszédszakaszokból felépülő hosszabb beszédészek, szövegek egészét tekintve nem mutatkoztott lassuló trend. Helyette a lassulások és a gyorsulások egyfajta kiegyenlítődése volt megfigyelhető. Ez nem azt jelentette, hogy ugyanolyan mértékű lassulás és gyorsulás volt megfigyelhető, vagy ugyanannyi gyorsulás és lassulás jellemezte ezeket a szövegeket. Sokkal inkább a lassulások és a gyorsulások gyakoriságának és mértékének fordított aránya jelent meg a szövegekben. A teljes szövegfelolvasást és a teljes spontán beszédet vizsgálva ugyanis azt a jellemző mintázatot kaptuk, hogy a növekvő lépések gyakoribbak, mint a csökkenő lépések, de a növekvő lépések átlagosan kisebbek, mint a csökkenő lépések (CV-szakaszok, beszédhangok, V-szakaszok esetében).

Ez a mintázat részben magyarázható a szünetek körül megvalósuló lassuló és gyorsuló lépésekkel. A szünetet átívelő lépések gyakran nagymértékű csökkenő lépések voltak, és a szünetek előtt tendenciaszerűen több lassuló lépés jelent meg. Így a szünet körülötte szakaszokban tendenciaszerűen látható egy ahhoz hasonló mintázat, mint amilyen a teljes szövegek egészére jellemző. A hosszú szakaszokra kapott mintázatoktól viszont eltér, hogy a lassuló lépések gyakoriságának növekedésével a lassulás mértéke is növekvő tendenciát mutatott. További eltérésnek tekinthető, hogy a lassuló lépések aránya a szünet előtti helyzetekben nem volt kimagaslóan magas, többségében a közlés belseji lépésekhez hasonlított. A szünet előtti pozíció ugyanakkor nem tekinthető ekvivalensnek a közlés vagy hangsor végi helyzettel, hiszen nem minden szünet esik egybe tagmondat- vagy megnyilatkozashatárral, illetve elképzelhető szünet nélküli lezá-

rás is. Ráadásul a lassításokat befolyásolhatja a fonológiai felépítettség is, hiszen a magánhangzós szakaszok és mássalhangzós szakaszok másfajta mintázatot mutattak hosszú szövegekben. A szünetnek a lassulásokra és gyorsulásokra gyakorolt hatásairól alkotott képünket tehát tovább árnyalhatná a fonológiai felépítettségnek, a szünet funkcióinak és a temporális változások összefüggéseinek vizsgálata.

A gyorsulásoknak és lassulásoknak az itt tárgyalt megközelítése csak egy kis szeletét képezi a lehetőségeknek. A vizsgált beszédrész nagyságának változtatásával valamelyest más aspektusai jelentek meg az időbeli szerveződésnek. Így feltételezhető, hogy az alapegység nagyságának variálása vagy a lépések hosszának, esetleg számának megváltoztatása további temporális összefüggések feltárásához vezethet.

Irodalom

- Abercombie, David 1967. *Elements of general phonetics*. Aldine, Chicago.
- Arvaniti, Amalia 2012. The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics* 40. 351–373.
- Boersma, Paul – Weenink, David 2013. *Praat: doing phonetics by computer [Computer program]. Version 5.3*. <http://www.praat.org>. (Letöltés ideje: 2013. júl. 9.)
- Dellwo, Volker 2009. Choosing the right rate normalization method for measurements of speech rhythm. In Schmid, Stephan – Schwarzenbach, Michael – Studer-Joho, Dieter (eds.): *La dimensione temporale del parlato*. Torriana. 13–32. http://www.zo-ra.uzh.ch/45236/1/Dellwo_AISV2010.pdf.
- Gósy Mária 2000. A beszédrítmus elemzésének egy lehetséges megközelítése. *Magyar Nyelvőr* 124. 273–287.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris, Budapest.
- Gósy Mária 2008. Magyar spontánbeszéd-adatbázis – BEA. *Beszédkutatás* 2008. 194–207.
- Gósy Mária – Beke András 2010. Magánhangzó-időtartamok a spontán beszédben. *Magyar Nyelvőr* 134. 140–165.
- Grabe, Esther – Low, Ee Ling 2002. Durational variability in speech and the rhythm class hypothesis. *Papers in Laboratory Phonology* 7. 515–546.
- Gyüre, Balázs – Bartos, Imre – Jánosi, Imre Miklós 2007. Nonlinear statistics of daily temperature fluctuations reproduced in a laboratory experiment. *Physical Review E* 76. 037301. http://www.karman.elte.hu/janosi/pdf_pub/PRE07temp-fluct-exp.pdf.
- Hardcastle, William J. – Laver, John – Gibbon, Fiona E. 2010. *The handbook of phonetic sciences*. Blackwell, Oxford.
- Kassai Ilona 1979. *Időtartam és kvantitás a magyar nyelvben*. Nyelvtudományi Értekezések 112. Akadémiai Kiadó, Budapest.
- Kohári, Anna 2013. Temporal patterns of segments and intervals in Hungarian language. In Mertens, Piet – Simon, Anne Catherine (eds.): *Proceedings of the Prosody-Discourse Interface Conference (IDP)*. Leuven. 51–57. http://www.ling.arts.kuleuven.be/franitalco/idp2013/papers/Mertens_Simon_2013_Proceedings_IDP2013.pdf.
- Kovács Magdolna 2002. *Tendenciák és szabályszerűségek a magánhangzó-időtartamok produkciójában és percepciójában*. Debreceni Egyetem Kossuth Egyetemi Kiadója, Debrecen.
- Ladefoged, Peter 1975. *A course in phonetics*. Harcourt Brace Jovanovich, New York.

- Lehiste, Ilse 1972. The timing of utterances and linguistic boundaries. *The Journal of the Acoustical Society of America* 51. 2018–2024.
- de Looze, Celine 2010. *ADoReVA and ADoTeVA, two PRAAT plugins for the automatic detection of register and tempo variations*. http://celinedelooze.com/2010DeLooze_Submission_SPSASSD.pdf.
- Magdics Klára 1966. A magyar beszédhangok időtartama. *Nyelvtudományi Közlemények* 68. 125–139.
- Markó Alexandra 2010. A prozódia szerepe a spontán beszéd tagolásában. *Beszédkutató 2010*. 82–99.
- Markó Alexandra 2012. A magyar hangsúly realizációinak és észlelésének összefüggése felolvasásban és spontán beszédben. In Markó Alexandra (szerk.): *Beszédtudomány. Az anyanyelv-elsajátítástól a zöngékezdési időig*. ELTE Bölcsészettudományi Kar–MTA Nyelvtudományi Intézet, Budapest. 277–303.
- Meyer, Ernst Alfred – Gombocz, Zoltán 1909. *Zur Phonetik der ungarischen Sprache*. Berlins Buchdruckerei, Uppsala.
- Olaszy Gábor 2006. *Hangidőtartamok és időszerkezeti elemek a magyar beszédben*. Akadémiai Kiadó, Budapest.
- Olaszy Gábor 2010. A beszéd időszerkezetének szabályalapú modellezése. In Németh Géza – Olaszy Gábor (szerk.): *A magyar beszéd. Beszédkutató, beszédtechnológia, beszédinformációs rendszerek*. Akadémiai Kiadó, Budapest. 449–455.
- Pike, Kenneth 1945. *The intonation of American English*. University of Michigan Press, Ann-Arbor.
- Ramus, Franck – Nespor, Marina – Mehler, Jacques 1999. Correlates of linguistic rhythm in the speech signal. *Cognition* 72. 1–28.
- Roach, Peter 1982. On the distinction between 'stress-timed' and 'syllable-timed' languages. In Crystal, David (ed.): *Linguistic controversies*. Edward Arnold, London. 73–79.
- van Santen, Jan 1992. Contextual effects on vowel duration. *Speech Communication* 11. 513–546.
- Schiel, Florian 1999. Automatic phonetic transcription of non-prompted speech. In Ohala, J. John – Hasegawa, Yoko – Ohala, Manjari – Granville, Daniel – Bailey, Ashlee C. (eds.): *Proceedings of the 14th International Congress of Phonetic Sciences*. University of California, San Francisco. 607–610.
- Schreiber, Thomas – Schmitz, Andreas 2000. Surrogate time series. *Physica D*. 142. 346.
- Tarnóczy Tamás 1974. A magánhangzók akusztikai vizsgálatának problémái. *Általános Nyelvészeti Tanulmányok* 10. 153–180.
- Tóth Bálint 2013. *Rejtett Markov-modell alapú gépi beszédértékelés*. Doktori értekezés. BME, Budapest. http://www.omikk.bme.hu/collections/phd/Villamosmernoki_es_Informatikai_Kar/2013/Toth_Balint_Pal/ertekezes.pdf.
- Turk, Alice E. – Shattuck-Hufnagel, Stefanie 2000. Word-boundary-related duration patterns in English. *Journal of Phonetics* 28. 397–440.
- Váradi Viola 2013. *A spontán beszéd szegmentálása produkciós és percepciós szempontból*. Doktori értekezés. ELTE, Budapest. <http://doktori.btk.elte.hu/lingv/varadivola/diss.pdf>.
- Váradi Viola – Beke András 2013. Az artikulációs tempó variabilitása felolvasásban. *Beszédkutató 2013*. 26–41.
- White, Laurence – Mády, Katalin 2008. The long and the short and the final: Phonological vowel length and prosodic timing in Hungarian. In Barbosa, Plinio A. – Madureira, Sandra – Reis, César (eds.): *Proceedings of 4th Speech Prosody Conference*. Campinas, Brazil. 363–366.

MONYEEK – MORFOLÓGIAILAG EGYÉRTELMIKÉSZÍTETT ÓVODAI NYELVI KORPUSZ

Mátyus Kinga – Orosz György

Bevezetés

A Magyar Óvodai Nyelvi Korpusz (MONYEEK) elsősorban a gyermeknyelvi variabilitás vizsgálatára jött létre, 62 4,5–5,5 éves budapesti óvodás gyermekkel készített, egyenként 20–30 perces interjúból áll. Az átirat hozzávetőleg 39 000 megnyilatkozásban 140 000 szót tartalmaz. A felvételek 2012 tavaszán készültek.

A magyar gyermeknyelv kutatását a kezdetektől esetleírások, naplószerű közlések segítik, amelyek lehetővé teszik több nyelvi szint vizsgálatát is (vö. pl. Kenyeres 1926; S. Meggyes 1971; Lengyel 1976, 1977; Molnár 1978; Gósy 1981). Ezek az írások rendszerint egy gyermek beszédfejlődését követik nyomon. A nagyobb, több adatközlővel dolgozó gyermeknyelvi gyűjtemények jobbra egy-egy adott nyelvi szegmensre fókuszálnak (vö. pl. Lengyel (1981) a szórendre, Gósy (1984) a hangtani és szótani fejlődésre, Sebestyén (2006) a tipikus fejlődésű gyerekek hangképzésére stb.). Az akár vizsgálatvezetővel, akár a családtagokkal (anyával) vagy kortársakkal folytatott beszélgetéseket rögzítő felvételek száma meglehetősen alacsony. Ezek közül nagyságukkal is kiemelkedőek Réger Zita gyűjtései, melyek több szempontból vizsgálják a magyar és cigány gyerekek nyelvfejlődését. Egy longitudinális, összesen 2 évig tartó kutatásában az anya-gyermek kommunikáció alakulását vizsgálva többszörösen hátrányos helyzetű és iskolázott anyák gyermekükkel való beszédét rögzítette (Réger 1985). Réger Zita felvételeinek digitalizálása az MTA Nyelvtudományi Intézetében folyik. A Miki nevű adatközlővel készült felvételek a CHILDES adatbázisban is megtalálhatók (<http://childes.psy.cmu.edu/browser/index.php?url=Other/Hungarian/Reger/>). Egy másik nagyszabású projekt, a GABI (Gyermeknyelvi Beszédatadátbázis és Információtár, Bóna et al. 2014), melynek távlati célja 900 gyermekkel (3–18 évesek) készített interjú felvétele (részletesen lásd a jelen kötetben). A beszélgetéseket rögzítő felvételek több nyelvi szint vizsgálatára is alkalmasak, és egyben nem tipikus nyelvi fejlődésű gyermekek nyelvhasználatának vizsgálatához is kiindulópontként szolgálhatnak.

A MONYEEK célja az, hogy elegendő mennyiségű anyagot adjon a 4,5–5,5 éves gyerekek nyelvhasználatának vizsgálatához, valamint a felnőtt, írott nyelvi korpuszokra kidolgozott számítógépes nyelvészeti módszerek adaptálásával új lehetőségeket nyújtson a sok adatot tartalmazó nagyobb gyermeknyelvi korpuszok kezelésében, elemzésében.

Írásunkban részletesen bemutatjuk a korpusz felépítését, az interjúk módszertanát. Ismertetjük az átirat morfológiai annotálása során használt automatikus módszereket. Bemutatjuk még a bővített morfoszintaktikai címkekészletet, mely az annotálás alapját képezte. Részletezzük az automatikus elemzés során felhasznált algoritmusokat, továbbá becslést adunk az automatikus egyértelműsítő módszer teljesítményére, melyhez a korpusz manuálisan javított részét használjuk.

A korpusz felépítése

Mivel a korpusz elsősorban a gyermeknyelvi variabilitás vizsgálatára készült, fontos szempont volt különböző szociokulturális háttérrel rendelkező gyermekekkel felvételeket készíteni. A Központi Statisztikai Hivatal által kiadott Budapest statisztikai évkönyve (Vida 2010) kerületenkénti bontásban közli a személyi jövedelemadó adatait. Mivel egyéb, részletesebb szociológiai összehasonlítás nem állt rendelkezésünkre, ezen adatok alapján választottuk ki, mely kerületek óvodáiban készüljenek a felvételek. A legmagasabb az egy főre jutó sja a II. és XII. kerületben, míg a legalacsonyabb a XXI. kerületben (Vida 2010: 75). Így a korpuszban a II. és a XII. kerület választott óvodái képezik a magasabb szociokulturális háttérű csoportot, míg a XXI. kerületiek az alacsonyabb szociokulturális háttérű csoportot. A fiúkkal és lányokkal készített felvételek aránya közel azonos. Összesen 62 gyerekkel készült interjúk, egyenként 20–30 perc hosszúságúak. Az egyes csoportokhoz tartozó interjúk megoszlását mutatja az 1. táblázat.

1. táblázat: A MONYEEK interjúinak megoszlása csoportonként

	Magasabb szociokulturális háttér	Alacsonyabb szociokulturális háttér
Fiúk	16	17
Lányok	15	14

Az interjúk megkezdése előtt tesztfelvételeket készítettünk. Kipróbáltuk, hogy az egyes feladatok hogyan működnek a választott korcsoportnál. Az interjúba már csak azok a feladatok kerültek be, amelyeket sikeresnek ítéltünk. A feladatok részben megegyeznek a Pap Mária és Pléh Csaba vizsgálatában használtakkal (Pap–Pléh 1972). A cél az volt, hogy egy felnőtt vizsgálatvezetővel folytatott beszélgetés során minél több, minél inkább összefüggő beszédanyagot gyűjtsünk a gyerekektől.

A MONYEEK korpusz csaknem 39 000 megnyilatkozásban 140 000 szót tartalmaz. A vizsgálat jellegéből és az adatközlők életkorából adódóan az adatok mintegy felét a vizsgálatvezetői adatok teszik ki, hiszen a kisebb gyermekekkel készített felvételek során a vizsgálatvezető szerepe sokkal nagyobb, sokkal többet beszél. Összehasonlításképpen, a teljes Budapesti Szociolingvisztikai Interjú (BUSZI-2) korpuszban a szavak száma 268 537, ebből a

terepmunkások 95 146 szót használtak, vagyis a terepmunkások beszéde teszi ki a korpusz 35,43%-át (a pontos adatok elérhetők a <http://corpus.nytud.hu/bu-szi/> webhelyen megtalálható keresőprogram segítségével). Szabó (2012: 53) korpuszában az általános és középiskolás diákokkal készített beszélgetéseiben a hetedikes diákoknál 42%, tizenegyedikes diákoknál 31% az interjúvezető beszédének aránya – szintén szavakban mérve.

Az interjúk átírásához a CHILDES (Child Language Data Exchange System) CHAT átíró és kódoló formátumát használtuk (<http://chilides.psy.cmu.edu/manuals/chat.pdf> l. MacWhinney 2000). Ez azzal az előnnyel jár, hogy egy kidolgozott, átgondolt átírási rendszert alkalmazhattunk, emellett a program számos funkciója megkönnyíti az adatok automatikus vizsgálatát (pl. szavak száma, leghosszabb megnyilatkozások, átlagos megnyilatkozáshossz = MLU: Mean Length of Utterance stb.), valamint ha a korpusz bekerül a CHILDES adatbázisba, az egységes formátum megkönnyíti a feldolgozhatóságot. Az átirat formai szabályai mellett a MONYEEK morfológiai annotációja is a CHILDES specifikációját (<http://chilides.psy.cmu.edu/manuals/clan.pdf>) követi. Az átlagos megnyilatkozáshossz, vagyis MLU számítása kiemelt szerepet kap, hiszen ez a gyermeknyelvi vizsgálatok egyik legfontosabb mérőszáma.

Az interjúk felépítése

Minden felvétel előtt írásos beleegyező nyilatkozatot kértünk a szülőktől, melyben engedélyezik, hogy gyermekükkel tudományos célból hangfelvétel készüljön. Emellett egy kérdőív kitöltését is kértük tőlük, amelyben a gyerekek szociokulturális háttéréről, nyelvfejlődésüket befolyásoló történekekről kértünk adatokat. A kérdőív elkészítéséhez az országos kompetenciamérés tanulói kérdőívét vettük alapul (<https://www.kir.hu/okmfit/>), különösen azokat a kérdéseket, amelyek esetében kimutatható összefüggés állt fenn a nyelvhasználattal.

Az interjúk szigorú módszertan szerint készültek. Minden helyszínen a csoporttól külön, egy külön helyiségben, ám az óvodán belül, nem hangszigetelt szobában vettük fel a beszélgetést. Erre azért volt szükség, mert a csoportban túlságosan nagy lett volna a háttérzaj, valamint a sok egyéb történet elvonta volna a gyermekek figyelmét a feladatokról. A külön helyiségek (pl. logopédiai foglalkozások színhelye) néhol ugyan meglehetősen zajosak voltak (pl. a szomszéd szobában zajló foglalkozások miatt), emiatt sok felvétel fonetikai vizsgálatra nem alkalmas, de ismerős helyek voltak a gyerekek számára, így nem jelentettek külön gátlást. A felvétel megkezdése előtt a vizsgálatvezető lehetőség szerint több időt töltött a csoportban, hogy a gyerekek egy kicsit megismerjék.

Minden interjú az adatok (név, életkor) rövid felvételével, illetve a ruhára tűzött mikrofonnal való ismerkedéssel ("Most azt játsszuk, hogy te egy titkosügynök vagy.") kezdődött. A feladatok mindig ugyanabban a sorrend-

ben követik egymást, és egyre kevésbé megkonstruált szövegek alkotását kívánják meg az adatközlőtől.

Az első feladat egy történet hallás utáni elmesélése képek segítségével. Itt a vizsgálatvezető képek alapján elmesél (felolvas) egy történetet (*Zsuzsi és az állatok*), majd a gyermekeknek ezt a történetet kell, szintén a képek alapján elmesélniük, elismételniük. A *Zsuzsi és az állatok* nagyon jó bevezető feladatnak bizonyult, mivel egyszerű, így szinte az összes gyermeknek sikerélményt jelentett, és jó felvezetése volt a második, bonyolultabb feladatnak.

A második feladat Mercer Mayer széles körben ismert és nyelvészeti kutatásokban gyakran használt képtörténetének (*Frog, where are you?*) elmesélése anélkül, hogy előtte azt a vizsgálatvezető elmondaná. Itt a felnőtt minta hiányában a feladat jóval önállóbb szövegalkotást kíván meg a gyerekektől, mint az első feladat. A „békamese” a legtöbb gyerek számára a legérdekesebb, legmotiválóbbról részét jelentette az interjúnak.

Ezután két további, rövidebb, 4-4 képből álló képtörténet elmesélése, valamint kiegészítése (Szerinted miért történt ez így? Mi történt vajon ezután?) következett, szintén felnőtt minta nélkül. Ezek a képtörténetek szintén átvezetésként szolgáltak a következő feladathoz, a játékszabályok elmeséléséhez.

Minden gyermeket megkértünk arra, hogy egy játék szabályait mesélje el úgy, hogy azt a vizsgálatvezető meg tudja tanulni. Itt már csak maga a játék szerepelt vezérfonalként.

A feladatok között, ahogy lehetőség adódott, sor került a gyerekek által (az interjú során vagy azon kívül) felvetett témák megbeszélésére is. Ezekben a témákban (pl. tűzoltóállomás látogatása, születésnap zsúr szervezése, új háziállat, kedvenc sport) a gyerekek motiváltak voltak, szívesen meséltek.

Az interjúk részét képezte még egy külön modulként a „szép beszédről” való beszélgetés, mely a nyelvi ideológiák kialakulásának, fejlődésének vizsgálatát segíti (Szabó–Mátyus 2013).

Az elkészült interjúkat a CHILDES szabályai szerint átírtuk, majd az átiratokat a hanganyaggal összevetettük.

A morfológiai annotáció

Az átiratok elkészítését követően a korpusz morfoszintaktikai annotálását végeztük el. Az alábbiakban először ismertetjük a munkánk során felhasznált eszközöket, majd részletezzük az elkészült sémát, ezt követően pedig összefoglaljuk a korpusz feldolgozásának lépéseit.

Felhasznált erőforrások

Írott szövegek elemzése során a morfológiai egyértelműsítés olyan alapvető nyelvtechnológiai feladat, mely magában foglalja a mondatokban szereplő egyes szavak morfoszintaktikai címkéinek és lemmáinak megállapítását. Ennek eredményére további feldolgozóeszközök épülnek, így az egyértelműsítő rendszer megbízható működése és magas pontossága alapvető fontosságú.

Ezen nyelvtchnológiai szoftverek a cél érdekében legtöbbször szabályalapú vagy gépi tanuláson alapuló algoritmusokat használnak.

Magyarra csak néhány olyan szabadon elérhető alkalmazás létezik, mely képes teljes morfológiai egyértelműsítésre. Egy ilyen a magyarlanc (Zsibrita et al. 2013), mely mondatok teljes szintaktikai elemzését végzi gépi tanulási algoritmusokkal. Ezzel szemben a PurePos (Orosz–Novák 2013) csak morfológiaiilag egyértelműsíti a mondatok tokenjeit. Bár mindkét szoftver szabadon használható, az utóbbi nyílt forráskódú, így működése nagymértékben testre szabható.

Az egyértelműsítés feladatához a felügyelt gépi tanulást használó algoritmusok ún. manuálisan egyértelműsített tanítóanyagot használnak. Jelenleg az egyetlen ilyen szabadon elérhető magyar nyelvű erőforrás a Szeged Korpusz (Csendes et al. 2004), mely mintegy 1 200 000 tokent tartalmaz.

Számtalanszor megmutatták (pl. Orosz – Novák 2013), hogy gazdag morfológiával rendelkező nyelvek esetén egy morfológiai elemző integrált használata nagymértékben képes javítani az egyértelműsítő folyamat eredményességén. Figyelembe véve a fenti megfontolásokat, munkánkhoz a HuMor (Novák 2003) morfológiai elemző kimenetét a PurePos egyértelműsítővel kombináltan használjuk. Az utóbbi tanításához a Szeged Korpusz egy automatikusan konvertált változatát alkalmaztuk.

A morfológiai annotáció menete és a létrehozott séma

Az átirat annotációjának elkészítéséhez a CHILDES útmutatásait követtük. Ennek megfelelően a megnyilatkozások minden egyes tokenjéhez rendeltünk morfológiai elemzést, kivéve az újrakezdést, fragmentumot és kivehetetlen szöveget jelző szimbólumokhoz. Az útmutatásoktól eltérően az ismétlések minden elemét annotáltuk, mégpedig úgy, hogy azok összetevői egyenként elemzésre kerültek. A helyes MLU-értékekhez az ismétléseket a továbbiakban egy tokenként kezeljük.

A CHAT útmutató számos ajánlást tartalmaz a morfológiai elemzések reprezentációjával kapcsolatban. Ebben többek között bemutatásra kerül egy formalizmus a szótövek, toldalékok és a szófaji kategóriák együttes tárolásához, mely az átirat készítéséhez használt alkalmazásban elérhető ún. MOR parancs kimenetéhez illeszkedik. Ez a modul képes elemezni többek között olyan nyelvek szavait, mint az angol, francia, német stb., de pl. magyar nyelv esetén már nem alkalmazható. Feladatunk egy olyan annotált korpusz készítése, amely grammatikailag könnyen összehasonlítható írott szövegekkel is, ezért a fenti ajánlásoktól részben eltérve szótó és morfoszintaktikai címke párosaként reprezentálva rögzítjük a morfológiai annotációt. Ehhez a sémához a HuMor elemző kódkészletét és lemmatizálási algoritmusát vettük alapul.

Az átiratok elemzéséhez felhasznált, írott szövegeken alapuló automatikus algoritmusok nem bizonyultak megfelelő teljesítményűnek, így azokat több ponton módosítottuk. Első lépésként a HuMor beállításait és lexikonját kellett adaptálni. Mivel célunk egy olyan korpusz létrehozása volt, mely lehetővé te-

szi a gyermeknyelv morfológiai és szintaktikai jellemzőinek vizsgálatát, szükségesnek találtuk a címkekészlet módosítását. Gyakori jelenség a gyermeknyelvben a kicsinyítő képzők fokozott jelenléte (pl. *anyuka, brekus*), melyek az írott szövegekben csak elvétve fordulnak elő. A további feldolgozó algoritmusok támogatásának céljából ezt a jelenséget külön jelöltük: a szavak szótöve a nem kicsinyített alak lett, míg az összetett címkében megjelent a képzőt szimbolizáló jelölés. A beszélt nyelvre jellemző a kitöltött szünetek használata (pl. *ööö, mmm*), amelyek címkézésére egy új kategóriát használtunk, amelyet hozzáadtunk a HuMor lehetséges elemzéseéhez. A Szeged Korpusz elemzéseitől eltérően annotáltuk még a *hú, húha, hūha* stb. szavakat. Bár az elemző csak indulatszó szófaji kategóriát javasolt ezekre a tokenekre, de használatukat tekintve úgy találtuk, hogy közelebb állnak a mondatzavak csoportjához, ezért az írott nyelvi sémától itt is eltértünk. A létrehozott korpusz újdonsága még, hogy a hangutánzó szavak csoportját használatuknak megfelelően indulatszóként, illetve ragozott alakjaikat igeiként elemeztük.

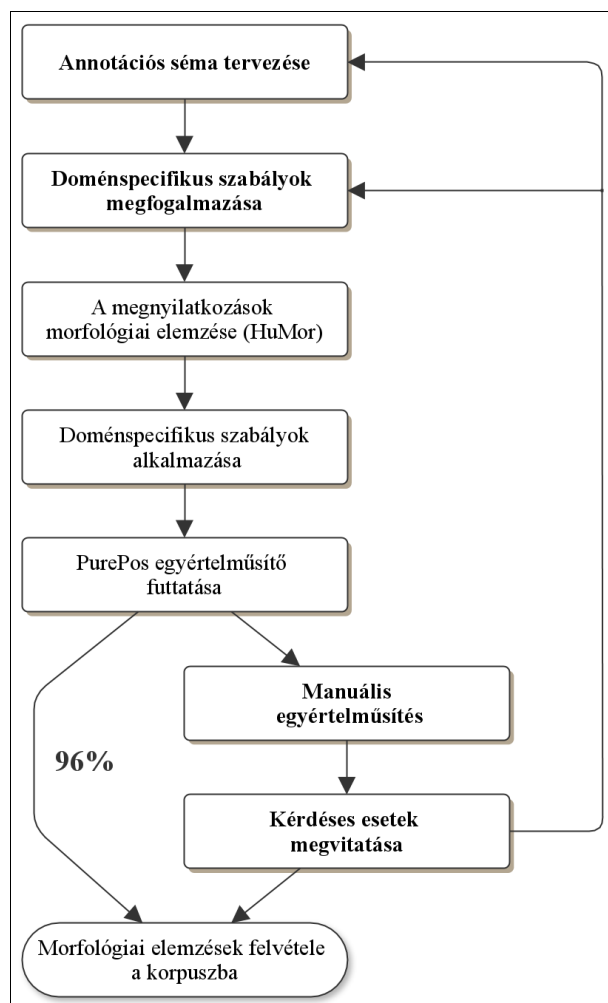
A beszélt nyelvben gyakoriak még az olyan szavak, melyek a beszélgetés folyamatosságát segítik elő, így önmagukban egy-egy megnyilatkozást alkotnak. Ilyenek például a *jó, rendben, ok, igazán, valóban* stb. szavak, melyeket a korpuszban mondatzó jelölést kaptak. Az annotálási folyamat automatikus módszerekkel való segítségéhez ezek az elemzések is bekerültek a HuMor lexikonjába. Mivel a beszélt nyelvben lexikalizálódott kifejezések, úgy mint a *légyszi, légyszíves*, a magyar helyesírástól eltérően egy tokenként kerültek átírásra, így ezek morfológiai címkéjét is meg kellett határozni. A szavak funkcióját vizsgálva úgy döntöttünk, hogy igei elemzést kapnak.

A bemutatott annotációs séma iteratív eljárás eredménye (1. ábra), mely automatikus eszközökön kívül emberi munkát is igényelt. Az alábbi lépéseken keresztül egyszerre fejlődött az elemzőlánc és a létrehozott séma:

1. A folyamat első lépéseként a rendelkezésre álló adatok alapján létrejött, illetve módosult az annotálási útmutató és ezzel együtt a korpusz elemzéséhez használt séma is.
2. A megfogalmazott iránymutatásoknak megfelelően változott a HuMor lexikonja, továbbá szabályalapú komponensekkel bővült az elemzési lánc.
3. A morfológiai elemző futtatását követően a doménspecifikus szabályok alkalmazásával szűrtük a lehetséges címkék és szótövek halmazát, amelyekből a PurePos egyértelműsítő választotta ki a helyesnek vélt elemzéseket.
4. Ezek után a manuális javítás következett, mely során a kérdéses eseteket az annotátorok megvitatták, és szükség szerint visszatértek az 1. vagy 2. lépéshez.
5. Végül automatikus eszközök segítségével a korpusz szavaihoz rögzítettük az elemzéseket.

Ez idáig 478 megnyilatkozást ellenőriztünk manuálisan, ennek eredményeképpen készült el a bemutatott annotálási séma és az ahhoz tartozó útmutató első verziója. Munkánk eredménye még, hogy a kezdeti alacsony pontos-

ságú feldolgozó algoritmusokat oly mértékben adaptáltuk, hogy ezek láncolata jelenleg mintegy 96%-os szószintű pontossággal képes egyértelműsített elemzések készítésére.



1. ábra

A korpuszhoz készült morfológiai annotáció létrehozásának lépései

További lépések

Célunk volt többek között olyan automatikus eljárások fejlesztése, melyek hatékonyan tudnak támogatni nyelvészeti kutatásokat, melyek a korpusz morfoszintaktikai viszonyaira irányulnak. Ennek érdekében szükséges a manuális annotálás folytatása, hogy az így létrejött etalont felhasználva az elemzőlánc pontosságát javítani lehessen. Távlati terveink közt szerepel még, hogy az átiratok egyes szekcióihoz (feladatokhoz) beszélőnként közreadjuk az automatikusan becsült, morfémában mért átlagos megnyilatkozáshosszértékeket (MLU) is. Ehhez első lépésként szeretnénk létrehozni egy automatikus megfelelőtétő eljárást, amely a korpusz annotációit egyértelműen leképezi a morfológiai elemző teljes elemzéseire. Erre épülve már megvalósítható az az algoritmus, mely a morfémák szintjén létrejött elemzésekre építve becsül minimum és maximum MLU-értékeket.

Irodalom

- Bóna Judit – Imre Angéla – Markó Alexandra – Váradi Viola – Gósy Mária 2014. GABI – gyermeknyelvi beszédadatbázis és információtár. *Beszédkutatás* 2014. 246–251.
- Csendes Dóra – Csirik János – Gyimóthy Tibor 2004. The Szeged Corpus: A POS tagged and syntactically annotated Hungarian natural language corpus. In Hansen-Schirra, Silvia – Oepen, Stephan – Uszkoreit, Hans (eds.): *Proceedings of the 5th International Workshop on Linguistically Interpreted Corpora (LINC 2004) at The 20th International Conference on Computational Linguistics (COLING 2004)*. University of Geneva, Geneva. 19–23.
- Gósy Mária 1981. A beszédhang kialakulása a gyermeknyelvben. *Magyar Fonetikai Füzetek* 7. 67–97.
- Gósy Mária 1984. *Hangtani és szótani vizsgálatok hároméves gyermekek nyelvén*. Nyelvtudományi Értekezések 119. Akadémiai Kiadó, Budapest.
- Kenyeres Elemér 1926. *A gyermek első szavai és a szófajok föllépése*. Kisdednevelés, Budapest.
- Lengyel Zsolt 1976. Formation and development of the verb system in the language of a two-year-old child. In: *Akten des I. Salzburger Kolloquium über Kindersprache*. Narr, Tübingen. 113–122.
- Lengyel Zsolt 1977. Developments of parts of speech in the Hungarian child language. *International Journal of Psycholinguistics*. 4. 50–64.
- Lengyel Zsolt 1981. *A gyermeknyelv*. Gondolat, Budapest.
- MacWhinney, Brian 2000. *The CHILDES Project: Tools for analyzing talk*. 3rd Edition. Lawrence Erlbaum Associates, Mahwah, NJ.
- S. Meggyes Klára 1971. *Egy kétéves gyermek nyelvi rendszere*. Nyelvtudományi Értekezések 73. Akadémiai Kiadó, Budapest.
- Molnár Ildikó 1978. A hanghelyettesítés típusai a gyermeknyelvben 18–21 hónapos kor között. *Magyar Fonetikai Füzetek* 1. 44–52.
- Novák Attila 2003. Milyen a Jó Humor? In: Alexin Zoltán – Csendes Dóra (szerk.): *Magyar Számítógépes Nyelvészeti Konferencia (MSZNY 2003)*. Szegedi Tudományegyetem, Szeged. 138–145.

- Orosz György – Novák Attila 2013. PurePos 2.0: a hybrid tool for morphological disambiguation. In Angelova, Galia – Bontcheva, Kalina – Mitkov, Ruslan (eds.): *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2013)*. Bulgaria, Hissar. 539–545.
- Pap Mária – Pléh Csaba 1972. Nyelvhasználat és társadalmi helyzet: Szociolingvisztikai vizsgálat hatéves fővárosi gyermekek körében. *Szociológia* 1/2. 211–235.
- Réger Zita 1985. Beszámoló a „Halmazottan hátrányos helyzet és a nyelvi fejlettség” című folyamatban lévő kutatási projektum néhány eredményéről. In *Műhelymunkák a nyelvészet és társtudományai köréből* I. 87–118.
- Sebestyén Tar Éva 2006. *A 3–6 éves kori fonológiai fejlődés mintázata a magyarban*. Open Art, Budapest.
- Szabó Tamás Péter 2012. „*Kirakunk táblákat, hogy csúnyán beszélni tilos*”: A javítás mint gyakorlat és mint téma diákok és tanáraik metanyelvében. Gramma, Duna-szerdahely/Dunajská Streda.
- Szabó Tamás Péter – Mátyus Kinga 2013. Óvodások és kisiskolások nyelvi ideológiái. In *XXIII. MANYE kongresszus előadásai*. Megjelenés alatt.
- Vida Judit (szerk.) 2010. *Budapest statisztikai évkönyve*. KSH, Budapest.
- Zsibrita János – Vincze Veronika – Farkas Richárd 2013. magyarlanc: A toolkit for morphological and dependency parsing of Hungarian. In Angelova, Galia – Bontcheva, Kalina – Mitkov, Ruslan (eds.): *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2013)*. Bulgaria, Hissar. 763–771.

A Magyar Óvodai Nyelvi Korpusz a CESAR (Central and South-East European Resources) projekt támogatásával jött létre.

A MONYEEK korpusz morfológiai annotációjának kidolgozása a TÁMOP-4.2.1.B-11/2/KMR-2011-0002 projekt támogatásával készült.

GABI – GYERMEKNYELVI BESZÉDADATBÁZIS ÉS INFORMÁCIÓTÁR

**Bóna Judit – Imre Angéla – Markó Alexandra –
Váradi Viola – Gósy Mária**

Gyermekegyes nyelvészeti beszédatadattárak

A nyelvészeti és a beszédkutatások egyre inkább igénylik, hogy az eredményeik nagy mintán is igazolhatók legyenek, illetve sok-sok adatközlő adatainak, nyelvi és beszédmintáinak az elemzésével jussanak el az általános következtetésekre. A beszédben ez különösen fontos, mivel a verbális megnyilatkozások nagy változatosságot mutatnak mind a beszélők között, mind a beszélőkön belül. Az univerzális jegyeket tehát csak „több beszélő elegendően nagy adatháttere” segítségével lehet megfelelően vizsgálni (Fant et al. 1991; Vicsi 2010: 261).

Amíg egyre több felnőtt nyelvi beszédatadattárist hoznak létre (vö. Gósy 2008), addig a nemzetközi szakirodalomban is csak viszonylag kevés gyermekbeszéd-adattár létezik – bár egyre nagyobb az igény a gyermekegyes nyelvadatbázisalapú vizsgálatára (Tomasello–Stahl 2004).

A legismertebb nemzetközi gyermekegyes nyelvészeti beszédatadattár a CHILDES – Child Language Data Exchange System (vö. MacWhinney–Snow 1990), amelyet különböző anyanyelvet elsajátító gyermekek beszéde alapján hoztak létre, és amely lehetőséget nyújt az univerzális jelenségek leírására, a szókincs és a grammatikai fejlődés sajátosságainak elemzésére. Az adattár szabadon elérhető és használható (<http://childes.psy.cmu.edu>), és lehetőség van újabb felvételek feltöltésére is. Magyar gyermekektől is tartalmaz anyagot. A CHILDES anyagán készült kutatásokból eddig több mint 3000 publikáció készült.

Több más (főként amerikai és brit) gyermekbeszéd-gyűjtemény létezik; például a Kids’ Audio Speech Corpus amerikai óvodás és iskolás gyermekek olvasott és spontánbeszéd-szövegeinek gyűjteménye (Eskenazi 1996); míg beszédtechnológiai célból, beszédfelismerőkhöz is hoztak létre gyermekbeszéd-korpuszokat (Shobaki et al. 2000; Kazemzadeh et al. 2005; D’Arcy et al. 2004). Többféle beszédhelyzetben (a narratívák mellett társalgási és mesemondási helyzetben), egynyelvű és kétnyelvű gyermekekkel is készültek felvételek a SALT adattárakban (<http://www.saltsoftware.com/salt/databases/#>). Léteznek továbbá orosz adattárak (INFANTRU és CHILDRU, Lyakso et al. 2010), illetve többnyelvű, akcentusos beszédet rögzítő gyermekbeszéd-adattárak is (Children News, vö. Raab et al. 2007).

Az első magyar nyelvű gyermekbeszéd-felvételeket a Hegedűs-archívumban találjuk (Menyhárt 2012). Ebben 31 falusi gyermek 125 percnyi hanganyaga található, a gyermekek életkora 10–16 év. Beszédtechnológiai alkalmazásokhoz készült a SPECO gyermekbeszéd-adatbázis, amely 5–10 éves gyermekek által bemondott (ismételt, ill. olvasott) szótagokat, szavakat, mondatokat tartalmaz (Csatári et al. 1999). A Magyar óvodás beszélt nyelvi korpusz (MONYEK) 4,5–5,5 éves gyermekek beszédprodukcióit tartalmazza többféle beszédhelyzetben (Mátyus–Orosz 2014). Metanyelvi és pszicholingvisztikai vizsgálatokra alkalmas Szabó (2012) korpusza, amelyben 1–4., 7. és 11. osztályos diákok (budapestiek, vidékiek és határon túliak) félig strukturált interjúi találhatók. Ezeket túl számos más célzott hangfelvétel készült különböző kutatási célokra óvodásokkal, kisiskolásokkal, tizenévesekkel (például Horváth 2006; Laczkó 2009; Neuberger 2013). Ugyanakkor továbbra is hiányzik egy széles életkori spektrumot átfogó, sok szempontú kutatásra is alkalmas, nagy mennyiségű hanganyagot tartalmazó gyermekbeszéd-adatbázis, amelyhez a kutatók hozzáférhetnek. Ezt a hiányt kívánja pótolni a következőkben bemutatásra kerülő GABI.

A GABI előzménye

Egy magyar nyelvű, sok beszélővel, széles életkori spektrumot lefedő gyermekbeszéd-adatbázis ötlete 2007-ben merült fel az ELTE Fonetikai Tanszékén, ekkor azonban pályázati támogatás hiányában nem valósultak meg a tervezett felvételek. Ugyanakkor a tanszék oktatóinak, doktoranduszainak és az alapképzésben részt vevő hallgatóinak is egyre nagyobb figyelmre fordult a gyermeknyelvi vizsgálatok felé. 2011-ben próbafelvételek készültek egy gyermekbeszéd-adatbázis tervének kidolgozásához (Imre 2011); míg végül 2013-ban gyermeknyelvi kutatócsoport alakult a tanszéken, amelynek egyik fő célja a beszédadatbázis létrehozása.

A GABI fejlesztése

A GABI (Gyermeknyelvi beszédAdatBázis és Információtár) egy fejlesztés alatt álló, fonetikailag megalapozott, többfunkciós spontánbeszéd-adatbázis, amelynek célja az, hogy a mai köznyelvet beszélő gyermekek beszédét rögzítse, továbbá anyagot biztosítson különféle kutatásokhoz.

Az adatbázis 3–18 éves gyermekek/fiatalok beszédprodukciójának felvételeit rögzíti, a távlati tervek szerint mindegyik életkorban 50–50 főét (a végső cél tehát mintegy 900 gyermek beszédfelvétele). Az adatbázisban elsősorban egynyelvű gyermekek beszédének rögzítését tervezzük, de próba jelleggel már elkészült 10 magyar–német két nyelvű gyermek hangfelvétele is.

Az adatbázis óriási jelentőségű, hiszen korábban még nem készült hasonló nagyságú, hasonlóan sokféle beszédfeladatot rögzítő beszédadatbázis gyermekekkel. Különösen kevés adatunk van a 9–10 éves kor fölötti gyermekek/fiatalok beszédéről, az adatbázisban velük is készülnek hangfelvételek.

Az adatbázis fejlesztése 2013 februárjában kezdődött az ELTE Fonetikai Tanszékén; a készítésében a gyermeknyelvi kutatócsoport tagjai vesznek részt, akik a tanszéki oktatók mellett kollégák, doktoranduszok, egyetemi hallgatók. A felvételi protokoll kialakításában számos, a gyermeknyelvi felvételek készítésében jártas kolléga véleményét meghallgattuk, konzultáltunk a BEszélt nyelvi Adatbázis (BEA) fejlesztőivel (Gósy et al. 2012), illetve próbafelvételeket készítettünk különböző életkorú gyermekekkel.

Egy gyermek felvétele mintegy 30-40 percet vesz igénybe. Jelenleg – alig egy év alatt, amióta a felvételi protokoll kidolgozását megkezdtük – 80, a protokoll szerint rögzített felvétel készült el.

A felvételi protokoll

A hanganyagok rögzítése a gyermekek természetes környezetében, az óvoda, iskola, esetleg saját otthon egy csendes termében történik. A felvételek elkészítése előtt ismertetjük a szülőkkel az adatbázis célját, a felvételek menetét, és írásbeli hozzájárulást kérünk tőlük a gyermekük beszédének rögzítéséhez. Emellett egy részletes anamnézist is kitöltetünk velük, ebben a gyermek születési körülményeire, szociális helyzetére és a beszédfejlődésére, illetve az azt befolyásoló tényezőkre kérdezzük rá. Az aláírt hozzájárulás és a kitöltött anamnézislap kézhezvétele után kezdjük el a felvételkészítést.

A felvételeket digitális formában rögzítjük 44,1 kHz-es mintavételezéssel. Tárolás: 16 bit, 86 kbit/s, mono. Bár nem stúdióban készülnek a felvételek, bizonyos fonetikai elemzésekre is alkalmasak. A gyermekek adatait titkosítjuk, és a felvételek rögzítése és feldolgozása során is szigorúan betartjuk a törvényi előírásokat. (A munkánk során alkalmazott etikai kódex a <http://spontanbeszed.hu/> oldalon olvasható.)

A gyermekek életkorától függően többféle változatban alakítottuk ki a protokollt. Két fő életkori csoportot alkottunk: 3–9 évesek; 9–18 évesek csoportja. (A 9 éves korosztályban mindkét típusú protokollal készülnek felvételek, azaz a 3–9 éves korúaknak készült protokollal és az idősebbeknek összeállított protokollal is. Így lehetővé válik, hogy ezt az életkori csoportot mindegyik feladatban összehasonlíthassuk mind a kicsi gyermekekkel, mind a kamaszodó és a középiskolás korosztállyal is.) Mindkét életkori csoport felvételi protokolljának kialakításához a BEA adatbázis protokollja volt az alap.

Protokoll a 3–9 éves kor közötti gyermekek felvételéhez

A gyermekekkel a következő beszédanyagot rögzítjük:

1. 15 különféle modalitású, hosszúságú, fonetikailag változatos mondat utánmondása. Például: *A gyermekek bukfencezni is megtanulnak testnevelés-órán. Nemsokára odaérünk, ugye? A gyöngyhalászok hosszú ideig képesek a víz alatt maradni.*

2. Spontán narratíva a gyermeket is érdeklő témákról (az óvodáról/iskoláról, a szabadidőről, közeli ünnepekről, eseményekről, utazásokról). Ha a

gyermek nem közlékeny, megkérjük, hogy mondja el egy csapatjáték játékszabályait stb.

3. 20 szó és kifejezés definiáltatása. Például: *lift, mérleg, szírom, lóvá tesz*.

4. Egy hallott szöveg tartalmának elmesélése. Ehhez a feladathoz a korcsoporton belül további életkori csoportokat alakítottunk ki: 3–4 évesek; 5–6 évesek; 7–9 évesek. Mindegyik csoportnak a saját életkorához mérten választottunk visszamondandó szöveget.

5. Mondatok felolvasása (2. osztályos kortól). A mondatok egy része megegyezik az 1. feladat mondataival, a másik része párbeszédet alkot. Például:

(1) Nincs kedved velem jönni?

Miért, hová megyünk?

Gondoltam, kimehetnénk a játszótérre.

6. Történetmesélés képsorozat alapján. A protokoll három különböző, egyenként hat képből álló mesés történetet tartalmaz.

Protokoll a 9–18 éves kor közötti gyermekek/fiatalok felvételéhez

A protokollt a BEA (Beszélt nyelvi adatbázis) alapján állítottuk össze.

1. 15 különféle modalitású, hosszúságú, fonetikailag változatos mondat utánmondása. A mondatok megegyeznek a fiatalabb gyermekeknek összeállított mondatokkal.

2. Spontán narratíva rögzítése.

3. 20 szó és kifejezés definiáltatása. A feladat megegyezik a fiatalabb gyermekeknek szánt definíciós feladattal.

4. Két hallott szöveg tartalmának elmesélése. A szövegek megegyeznek a BEA-ban használt szövegekkel (vö. Gósy et al. 2012).

5. Szöveg- és mondatfelolvasás. (A szöveg megegyezik a BEA-ban használt szöveggel.)

6. Történetmesélés képsorozat alapján. A képsorozatok megegyeznek a kisebb gyermekeknek szánt képsorozatokkal.

7. Vita két gyermek között egy, a felvételvezető által megadott témáról. Például: *Jó-e az ötfokozatú értékelés? Jó-e az iskolai egyenruha? Az autó helyett vissza kellene térni a kerékpárhoz.*

A GABI lejegyzése

A felvételek lejegyzése megkezdődött. A lejegyzéshez a BEA adatbázisban alkalmazott elveket használjuk (Gyarmathy–Neuberger 2011; Gósy et al. 2012), a hangfájlok annotálását jelenleg a Transcriber szoftverrel .trs-ben készítjük el. A későbbiekben tervezzük, hogy ezeket a fájlokat ellenőrzés után .TextGrid (Praat) formátumúvá alakítjuk.

Távlati tervek, kutatások

A GABI anyagán már most is megkezdődtek a kutatások, több konferencia-előadás hangzott el az adatbázison készült vizsgálatok eredményeiből.

Ezek egyelőre a szavak definíciójának sajátosságaira koncentráltak (Imre–Hertendi 2013; Nagy-Varga 2013). Több szakdolgozat készül jelenleg a gyermekek beszédproduktóinak temporális elemzéséből és a megakadásjelenségek vizsgálatából is.

Az adatbázis fejlesztése folyamatosan zajlik, az aktuális állapota nyomon követhető a www.fonetikaitanszek.hu honlapon.

Irodalom

- Csatári, Ferenc – Bakcsi, Zsolt – Vicsi, Klára 1999. A Hungarian child database for speech processing applications. In: *ESCA Proceedings*. BME, Budapest. 2231–2234.
- D'Arcy, Shona M. – Wong, Lit Ping – Russel, Martin J. 2004. Recognition of read and spontaneous children's speech using two new corpora. In: *Proceedings of ICSLP*. 1473–1476.
- Eskenazi, Maxine S. 1996. KIDS: a database of children's speech. *The Journal of the Acoustical Society of America* 100. 2759.
- Fant, Gunnar – Kruckenberg, Anita – Nord, Lennart 1991. Prosodic and segmental speaker variations. *Speech Communication* 10. 521–531.
- Gósy Mária 2008. Magyar spontánbeszéd-adatbázis – BEA. *Beszédkutatás* 2008. 194–207.
- Gósy Mária – Gyarmathy Dorottya – Horváth Viktória – Grácz Tekla Etelka – Beke András – Neuberger Tilda – Nikléczy Péter 2012. BEA: Beszélt nyelvi adatbázis. In Gósy Mária (szerk.): *Beszéd, adatbázis, kutatások*. Akadémiai Kiadó, Budapest, 9–24.
- Gyarmathy Dorottya – Neuberger Tilda 2011. A BEA-adatbázis alkalmazásfüggő lejegyzései. *Beszédkutatás* 2011. 109–120.
- Horváth Viktória 2006. A spontán beszéd és a beszédfeldolgozás összefüggései gyerekeknél. *Beszédkutatás* 2006. 134–146.
- Imre Angéla 2011. *Mit árulnak el a gyermeknyelvi felvételek?* Előadás a Beszédkutatás 2011 konferencián. Budapest, 2011. október 27–28.
- Imre Angéla – Hertendi Márta 2013. *A beszédtervezés stratégiai különböző életkorokban*. Előadás a Pszicholingvisztikai Nyári Egyetemen. Balatonalmádi, 2013. május 26–30.
- Kazemzadeh, Abe – You, Hong – Iseli, Markus – Jones, Barbara – Cui, Xiaodong – Heritage, Margaret – Price, Patti – Anderson, Elaine – Narayanan, Shrikanth – Alwan, Abeer 2005. TBALL data collection: The making of a young children's speech corpus. In: *INTERSPEECH 2005*. 1581–1584.
- Laczkó Mária 2009. Tizenévesek beszédének fonetikai és stilisztikai elemzése. *Anyanyelv-pedagógia* 2009/1. <http://www.anyanyelv-pedagogia.hu/cikkek.php?id=151>. (Letöltés ideje: 2013. 10. 29.)
- Lyakso, Elena E. – Frolova, Olga V. – Kurazhova, Anna V. – Gaikova, Julia S. 2010. Russian infants and children's sounds and speech corpora for language acquisition studies. In *INTERSPEECH 2010*, 1878–1881.
- MacWhinney, Brian – Snow, Catherine 1990. The Child Language Data Exchange System: An update. *Journal of Child Language* 17. 457–472.

- Mátyus Kinga – Orosz György 2014. MONYEEK – Morfológiaiilag egyértelműsített óvodai nyelvi korpusz. *Beszéd kutatás* 2014. 237–245.
- Menyhárt Krisztina 2012. A beszéd temporális jellemzői 60 évvel ezelőtti gyermek beszélőknél. *Beszéd kutatás* 2012. 246–259.
- Nagy-Varga Zsolt 2013. *Definícióalkotási stratégiák tizenéves diákok körében*. Előadás a *Beszéd kutatás* 2013 konferencián. Budapest, 2013. november 14–15.
- Neuberger Tilda 2013. A spontán beszéd temporális sajátosságai 6–14 év közötti gyermekeknél. *Anyanyelv-pedagógia* 2013/2. <http://www.anyanyelv-pedagogia.hu/cikkek.php?id=451>. (Letöltés ideje: 2013. 10. 29.)
- Raab, Martin – Gruhn, Rainer – Noeth, Elmar 2007. Non-native speech databases. In *IEEE Workshop on Automatic Speech Recognition & Understanding, ASRU*. 413–418.
- Shobaki, Khaldoun – Hosom, John-Paul – Cole, Ronald 2000. The OGI kids' speech corpus and recognizers. In: *Proceedings of ICSLP-2000*, Beijing, China. 258–261.
- Szabó Tamás Péter 2012. „Kirakunk táblákat, hogy csúnyán beszélni tilos”. *A javítás mint gyakorlat és mint téma diákok és tanáraik metanyelvében*. Gramma, Duna-szerdahely.
- Tomasello, Michael – Stahl, Daniel 2004. Sampling children's spontaneous speech: How much is enough? *Journal of Child Language* 31. 101–121.
- Vicsi Klára 2010. Adatbázisok a beszédtechnológia szolgálatában. In: Németh Géza – Olasz György (szerk.): *A magyar beszéd. Beszéd kutatás, beszédtechnológia, beszédinformációs rendszerek*. Akadémiai Kiadó, Budapest. 261–268.

A szerzők köszönetet mondanak a következő személyeknek, akik részt vettek az adatbázis fejlesztésében: Auszmann Anita, Bereznai Anja, Dobay József, Domokos Ildikó, Hertendi Márta, Jankovics Julianna, Jordanidis Ágnes, Katona Krisztina, Kővári Bettina, Krepsz Valéria, Lengyel Boglárka, Nagy-Varga Zsolt, Németh Máttyás, Neuberger Tilda, Pantóné Naszályi Dóra, Szabó Ágnes, Szabó Beatrix Xénia, Vakula Tímea, Varjasi Szabolcs.

„NYELVBOTLÁS”-KORPUSZ

10. rész

Az első magyar valós idejű, jegyzeteléses technikával gyűjtött megakadás-jelenség-korpuszt 2004 óta adjuk közre folyóiratunkban:

Beszéd kutatás 2004. 19–186.: 5139 adat;

Beszéd kutatás 2005. 145–173.: 761 adat;

Beszéd kutatás 2006. 231–247.: 388 adat;

Beszéd kutatás 2007. 187–198.: 244 adat;

Beszéd kutatás 2008. 221–239.: 444 adat;

Beszéd kutatás 2009. 257–267.: 176 adat;

Beszéd kutatás 2010. 283–291.: 179 adat;

Beszéd kutatás 2011. 149–165.: 429 adat;

Beszéd kutatás 2012. 301–313.: 298 adat.

Eddig tehát összesen 8058 megakadásjelenséget elemeztünk, ebben a részben pedig újabb 602 tételt adunk közre.

Értelemszerűen ez a korpusz a hiba típusúak gyűjteménye. A kezdetektől számos, különféle szempontú tanulmány, szakdolgozat, disszertáció készült ebben a témakörben. Mindez a tapasztalat oda vezetett, hogy a „Nyelvbotlás”-korpusz osztályozási kategóriáit 2009-től egyszerűsítsük. Az alapvető ok az volt, hogy a korpuszt használóknak nagyobb szabadságot biztosítsunk egy adott jelenség mélyreható elemzéséhez.

Az egyszerűsített osztályozás elve az, hogy minden esetben a felszínen tapasztalható jelenséget soroljuk kategóriákba, a finomelemzések az adott jelenség kutatójára várnak. A fő kategóriák és meghatározásuk a következők:

1. Téves szó

A megfelelő szó helyett egy másik szó megjelenése a felszínen. Ez a kategória tartalmazza a freudi elszólásokat és a malapropizmusokat is, amelyek további elemzésekkel különíthetők el.

2. Grammatikai hiba

A köznyelvi normának ellentmondó morfológiai/szintaktikai struktúra.

3. Kontamináció

Két nyelvi jel (szó, szerkezet) összeolvadása, vegyülése.

4. Téves kezdés

Beszédhang vagy beszédhangsorozat ejtése, amely azonban nem szó.

5. Nyelvem hegyén van jelenség

A szándékolt szó jelentésének és morfológiai struktúrájának ismerete, de a fonológiai/fonetikai szerkezet kiejtésének (részleges) gátoltsága.

6. Perszeveráció

Egy kiejtett nyelvi elem szándéktalan, ismételt megjelenése a közlés egy későbbi időpontjában.

7. Anticipáció

Egy szándékozott nyelvi elem megjelenése a közlés egy korábbi időpontjában is.

8. Metatézis

Nyelvi elemek sorrendjének felcserélődése.

9. Egyszerű nyelvbotlás

A beszédtervezés artikulációs tervezési szintjén, ill. az artikulációban létrejött hiba, amely semmilyen egyéb okkal nem magyarázható. Altípusai: a) betoldás, b) helyettesítés, c) kiesés

10. Több típusba sorolható jelenségek

Azok az adatok kerülnek ebbe a kategóriákba, amelyek létrejöttében a felszíni elemzés alapján több rejtetten működő folyamat hibája is feltételezhető.

1. Téves szó

Megakadás	Szándékolt közlés
A cukkinit lehet belerendelni dehogya rendelni reszelni.	
A fejemre teszem ezt a hallót □ fejhallgatót.	
A felvételi feltétele, hogy a gyerek halmazosan hátrányos helyzetű legyen.	halmozottan
A határos □ hatályos törvények szerint...	
A házasság élet zavarba kelt □ hoz.	
A Kheopsz fáraó fáraó na piramisról ír benne.	
A lottó ehető számai...	e heti
A MASZAT az az első magyar űrhord □ műhold.	
A nagybajuszú tündérek sétálgatni mentek.	cincérek
A petrezselyem zöltség is kell.	zöldje
A szupraszegmentumokat sokkal részletesebben dolgozták ki percepció □ produkció szempontból.	
A te mikroszkóp □ Microsoft-fiókkal jelentkezünk be.	
Akkor nem tudok nekik adni □ vinni cuccot.	
Alakul a dolog az elmúlt □ következő hatnapban.	
Amúgy meg köleses □ köménymagos.	
Az egyedüli dolog, amire a Vidámparkban felmernék ülni, az a forgókerék.	óriáskerék

Megakadás	Szándékolt közlés
Az ilyen profiloknak mint te, nincs szüksége erre.	profiknak
Az ógrammatikusok úgy gondolták, hogy a nyelv egy élő orgazmus öö organizmus!	
Az újabb nyelvtanok már egy kicsit óvatosabban foglalkoznak, fogalmaznak.	
Azért picit most helyre kell, hogy álljon a pólus vagy öö pulzus.	
Azért szerintem egy akkora lakást lehet venni tizenegy-két euróért □ millióért.	
Azt tudom magyarázatként, vagyis azt gondolom magyarázatként...	
Balázs hozott egy üveg □ pohár bort.	
Bár a nyelvi képességeik jók, de matematikából kiválóan teljesítenek.	a nyelvi képességeik rosszak
Bárki bármennyire □ bármilyen típusú hadaró.	
Baromi stresszes az életem, mostanában minden automatikusan váratlanul történik.	állandóan
Céljai elismerése érdekében.	elérése
Csak mi megyünk a vo □ naton?	buszon
Csak nem a Láncember.	Hídember
Egyik nagy kedvenc kajám a kenőmájas kolbásszal és mustárral. Vagy mi. Majonézzel.	
Egy-két helyen még már pirul.	
Eh, ennyi munka mellett nem is lehet dolgozni!	pihenni
Ekkora mogyoró diófánk van □ vagy mi.	
Elengedi a feje mellett öö a füle mögött öö a füle mellett!	
...és akkor így elhalkul hal □ halványul a képernyő.	
Ez a program a Balaton partján rajzik.	zajlik
Ez egy nagyon jó eszmecsere.	gondolat
Ez olyan, mint a jedik fémkardja.	fénykardja
Ez volt a hab a tésztán öö min is? Tortán!	
Ezek a zöld □ vagy mi piros bútorok.	
Ezért is gyakoriak, illetve ritkák a téves szótlálások.	
Felvettünk egy videót □ videót, interjút!	
Gondolkodnia kell az ő utánképzéséről □ továbbképzéséről.	

Megakadás	Szándékolt közlés
Ha nem tudtok nyugton enni, akkor menjetek vacsorázni!	menjetek fürödni
Hány zsömlét öö kiflit hozzak?	
Határozatlan volt, mint a határozatlan rago-	ragozás
dás.	
Hogy lehet ilyen szép egy panda □ jegesmaci!	
Holnap akkor pörkölt krumplit csinállok, jó?	
Vagyis nem, mert paprikás krumplit.	
Holnap beviszem, elmosatom □ megmosatom az autót.	
...illetve évente kontroll □ on, illetve kontrollvizsgálaton megjelenni...	
Ilyen túlnyomás van a palackban, és ha kinyitod, kijön belőle a Fornetti.	konfetti
Ismét drágább lesz a parkolás □ tankolás.	
Itt a tilos, hol a piros?	Itt a piros, hol a piros?
Itt nem régen van egy babzsákbolt öö nem messze.	
Ja, én mostanában mindig csirkebogyóteát iszok kajálásnál, mert még mindig fáj kicsit a torkom.	csipkebogyóteát
Jaj, ez a kutya olyan rokolyás. □ Izé, rigolyás.	
Kerülni kell a túlzott körülmények □ követelmények támasztását.	
Két dekásat tudtam szerezni □ két deciset.	
Kettő egész kettő tized óra □ század!	
Kezdek egyből □ újból.	
Körzetek szerint van, még csak nem is állatok szerint öö államok.	
...közép-obi nyelvállás öö nyelvjárás.	
Küldjön egész oldalas képet □ egész alakosat.	
Lacinak mondhatod, hogy boldog □ jó étvágyat.	
Látod, milyen illata van? Vagyis érzed?	
Láttátok a Miami Beach-et?	Mamma miát
Légy szíves, öltözködj gyorsan, és gyere!	mosakodj
Majd a DM-ben veszünk láthatatlan körömlakkot, dehogy láthatatlant, átlátszót.	
Majd meg kell nézni a naplót □ nem, a naptárat.	

Megakadás	Szándékolt közlés
Majd mondd azt, hogy ismert öö ismeretlen emberek között gyűjtötted.	
Májusban kaptam azt a bor □ azt a viszkit.	
Már így rajtam □ hozzám képest...	
Már kezdek enni öö éhes lenni.	
Már megint mi szidunk.	szívunk
Már nem lakik otthon, mert férjhez □ feleségül ment □ megházasodott.	
Másfél perc nem túl elég □ nem túl sok rá.	
Melyikük Kempelen öö Farkas Bertalan?	
Menjünk ebédelni az aluljáróba.	alagsorba
Mi vagyunk a legsörösebb sátán □ öö sátor	
Milyen intenzitással kell ezt a hadarást □ terápiát végezni?	
Milyen ízű füstölőt szeretnél? Van mézes, vanília...	milyen illatú
Minden szentnek maga felé folyik a keze	hajlik
Mindenki vegye elő a zöld nagyságú könyvét!	színű
Mindjárt megtömöm a mandulát és, mehetünk.	megtöröm
Most akkor a ti kutyátokban is van már ilyen csipsz?	chip
Nagyon jó ez a függöny, mert reggel nem süttött a fejembe □ szemembe a nap.	
Nagyon jó órát lehet ebből tanítani öö tartani.	
Nem adnak időseknek szóló sorozato öö műsorokat.	
Nem azon a négy órán □ vagyis negyed órán múlik.	
Nem hiszem el, hogy ez van, tényleg ezért tanultam fel hajnali 5-kor?	keltem fel
Nem kellene bika bika szarvashibákat benne hagyni.	
Nem minden színésznővel van probléma öö színésszel.	
Nem önteném betonba vagy öö ólomba, amit mondtam.	
Nem vitte rá a lényeg.	lélek
Néma kutyának anyja sem érti a szavát.	gyereknek
Odáig menjünk be a zebráig öö bójáig [a bója piros-fehér csíkos volt].	
Ortodox modellre tért át.	hitre

Megakadás	Szándékolt közlés
Ott jön egy busz öö kerékpár.	
Ő az ELTE óvodaképzőjére jár.	óvónőképzőjére
Persze, mert milliányi teljesen fölösleges indok között volt elrejtve. Olyan milliányi alkalmazás között, amely teljesen indokolatlanul tölti be a kezdőlapot.	fölösleges ikon
Rámcsukódik a villamos □ vagy vonatajtó.	
...rongyba csavart citrom öö téglá...	
Rózsaszín szívószálat köptem.	kaptam
Sima epres □ ecetes uborka.	
Sosem iszom vizet. Vagyis viszkit.	
Szerintem egy szál szegfűszeget. Szegfűszeget? Szegfűt.	
Szóval akkor mennyi három meg hét? Baszszus! Szóval mennyi három meg négy?	
Szúrjál bele egy ollót [a krupliba]. Egy villát.	
Tudtam, hogy ott nincs nálam térkép □ nem, szótár.	
Ültem a Golden Eye-on.	London Eye-on
Varrok neked egy talányt.	talárt
Vedd már elő a jegyeket, mindjárt itt a sofőr!	kalauz
Vettek egy transzformerszt.	Transportert
Vicces, hogy milyen ferde a fal. Vagy mi.	
Padló.	

2. Grammatikai hiba

Megakadás	Szándékolt közlés
A Ferencváros sikerült □ a Ferencvárosnak sikerült először győzedelmeskedni.	
A játék nyelvétől független az angol □ függetlenül az angol nyelv dominál.	
A jövő héten fog, aki menni?	lesz, aki megy
A megakadásom öö □ megakadások gyakran előfordulnak	
A mondatok ne legyen □ ne legyenek túl hosszúak.	
A óraszámok megnöttek.	az
A szókincsem is sokkal több.	bővebb
A vége felé elég nehéz feladatokkal szembesülnek a gyerek.	gyerekek

Megakadás	Szándékolt közlés
Amikor tegnap hallottuk a buszon azról... arról jutott eszedbe?	
Az ajtószekrényt nyitódott.	Az ajtószekrény nyitódott.
Az igazságértéket az utolsó elhangzott szó után lehessen □ lehetett megmondani.	
Azért az szép teljesítmény a könyves blogból □ blogtól!	
Azt a titkot el kell árulnom a tisztelt közön- séggel, hogy Móni három nappal ezelőtt még beszélni sem tudott.	közönségnek
Azt azért megemlítenénk öö megemlíteném	
Beszédhibások gyermekeknek öö □ be- szédhibás gyermekeknek	
Biztosan sikerülni fogsz □ fog.	színt
Csinálsz bele színet akkor.	keveset
Decemberben kevészet fogok dolgozni.	mosolyogsz
Egyfolytában engem nézel, de nem mosolygol egyfolytában.	
Ennél sokkal fontos, hogy tényleg fantasz- tikus produkciókat láthattunk.	Ennél sokkal fontosabb, hogy tényleg fantasztikus produkció- kat láthattunk.
Érzem, ahogy fickándoznak a nyelvemben nyelvemen az ízek.	
Ez egy olyan utópiatörténet, amelyben Éva megölje Ádámot.	megöli
Ez erre szól ez az allegória?	ennek szól
Ezek nem egy ilyen megfordítható állítá- sok.	nem ilyen megfordítható állítá- sok
Fogták a fejem a fejüket.	
Ha fáj a hasad, ráteszem a kezedet.	kezemet
ha körbenézitek.	körbenéztek
Hát ebben egészen biztosan vagyok.	biztos vagyok
Hatvanba szaladtam le érte a Tescoba □ hoz.	
Hogyan alakulnak a szavak száma?	alakul
Holnap csinálsz meg a számításokat!	csináld meg
Így is hallottunk □ hallottuk.	
Iszonyúan haragszik egymásra	haragszanak
Jobban megismerkedtem a kamerával, és nem félek már annyitól, mint májusban.	annyira
Kérdés az, hogy a statisztikailag is igazol- ható-e.	hogy statisztikailag is igazolha- tó-e

Megakadás	Szándékolt közlés
Kétféle módszerrel dolgoztam, egyrészt korpuszelemzésre <input type="checkbox"/> korpuszelemzés-sel...	
Legalább látszik, hogy te van ott.	te vagy ott
Megnéztem én se.	én is.
Megtaláljuk a céljunk <input type="checkbox"/> célunkat.	
Mem arról félek, hogy...	attól
Mem vagyunk biztos abban, hogy...	biztosak
Mi történnek a ragjaimmal?	Mi történt a ragjaimmal?
Mielőtt aggódtál, barátod is van.	aggódnál
Mit fog adni neked <input type="checkbox"/> nekem mamád?	
Módosítások van benne.	vannak
Most komolyan, tényleg a Nóri itt le?	ivott le
Most már szabadba <input type="checkbox"/> szabadon <input type="checkbox"/> szabad-jára engedi.	
Na, gyönyörű, nem igaz, hogy ennyiszer botlom meg a nyelvem.	botlik meg
Ne kimész a vonalból	menj ki
Nem elég a fizikai szemünket látni öö szemünkkel látni...	
Nem fogok túlaludni magam.	fogom
Nem tudom, miért rövidet akarnak iszni.	inni.
Nincs terjedelmi kötöttsége is.	sem
Oda kell figyelni vele, de én már megtanultam.	rá
Olaszul coriandoli. – Na, azt már elhisszük sszük.	
Olyan nyelvek, amiknél a szótagok...	amiknél
Ő sem fogja <input type="checkbox"/> sem fog oda bemenni.	
Ráadásul olyan illat, amit neked borzasztóan zavaró.	ami
Remélem, nem vannak fönt a csúcson.	nincsenek
Szerencséd lesz <input type="checkbox"/> lett.	
Szeretnék foglalkozni a témában.	témával
Szeretnék itt szakképesítést is szerezni az újságírásból <input type="checkbox"/> ról.	
Szóljatok, ha fáznak!	Szóljatok, ha fáztok!
Tehát nem konkrétan egy szavak <input type="checkbox"/> szó volt olvasható ott.	
Teljesen elzárkóznak a velünk való kommunikációval.	kommunikációtól
Több ezer létszámot bocsátottak el.	ezres

Megakadás

Többet nem viszünk magaddal.
 Többoldalas tárgyalások lesznek.
 Töltelekszavak is a csendesben lesznek többek.
 Úgy fog elindul ☐ beindulni a szülés...
 Vért izzadt ☐ izzadt.

Szándékolt közlés

magunkkal
 többoldalú
 A töltelekszavakból is a csendesben lesz több.

3. Kontamináció**Megakadás**

A frenébe!
 A srác lesütötte a fejét, meg se szólalt.

 Az a durva, hogy egy ilyen hajón 4-5 fős a legénység.
 Az adott beszély kitűzött szándéka...
 Az évek solyamár ☐ az évek folyamán.
 Az petrezselyemmorzsával készül ☐ zsemmorzsával, ez meg petrezselyemmel.
 Bolond lyukból cincognak az egerek.

 ...de a jelenleg Budapesten élek.
 De akkor még nem produkáltak senki.

 Eláll a szája.

 Én nem bírnám mint történelmesz a munkát.
 És mindezt az ő szeme hallatára sikerült végrehajtanotok?
 Ez a leggyorsabb állat a világon, ez a konkrét geopárd ☐ geopárd?! ☐ Gepárd
 Ez az érték irasztó borzasztó alacsony.
 Ez is emelte az est fénypontját.

 Ez nagyon jó, mert a kerodon tudod tartani.
 Ezentúl ha legyengül az immunszervezeted...
 Ezt még át kell ismételnünk, úgy tűnöm.
 Ezt próbálta Judit nekünk itt rávezetni...

Szándékolt közlés

fenébe × francba
 lesütötte a szemét × lehajtotta a fejét
 legénység x személyzet

 beszélő × személy
 során × folyamán

 Bolond lyukból bolond szél fúj
 × Nincs otthon a macska, cincognak az egerek.
 a jelen pillanatban × jelenleg
 produkált senki × produkáltak semmit
 eláll a lélegzete × tátva marad a szája
 történelem × történész

 füle hallatára × szeme láttára

 gepárd × leopárd

 irtó × borzasztó
 emelte az est fényét × az est fénypontja
 kezeden × karodon
 immunrendszered × szervezeted

 úgy tűnik × úgy gondolom
 ezt próbálta nekünk elmagyarázni × erre próbált minket rávezetni

Megakadás	Szándékolt közlés
Ha esik, ha puffan...	Ahogy esik, úgy puffan. × Ha esik, ha fúj.
Hadd verje ki őket az Isten hidege.	verje meg őket az Isten × megvesz az Isten hidege × kirázza a hideg
Hollan kezdünk?	honnan × hol
Így jár, aki pórul.	Így jár, aki... × pórul jár
Kend meg a szendvicseket sajjal!	sajt × vaj
Kérek egy kohár sört! Ja nem, egy porsóval!	pohár; korsóval!
...komorú képet fest...	szomorú × komor
Leintem, mint a taxiórát.	Leintem a taxit. × Leütöm, mint a taxiórát.
Mabi néni	Babi néni × Mama
Meglátjuk, ha az időjárás abbamarad öö a szél és az eső abbamarad.	a szél és az eső abbamarad × az időjárás megjavul
Mindenre csak húzza a fogát!	húzza a száját × szívja a fogát
Mindent megteszünk a szügyért □ na, szent ügyért.	
Ne hányd az ördögöt a falra!	ne fessd az ördögöt a falra × falra hányt borsó
Ne kísérsd az ördögöt a falra festeni!	Ne kísérsd az ördögöt! × Ne fessd az ördögöt a falra!
Ne már, összerontod az egészet!	összekevered × elrontod
Nem hagyom elengedni .	nem hagyom elmenni × nem akarom elengedni
Nem kürtölöm senkinek az orrára, ez a te magánéleted.	nem kürtölöm szét × nem kötöm senkinek az orrára
Nem tudom, nem rajtam függ.	rajtam múlik × tőlem függ
Nincs véletlenül moláv □ moláv...	MÁV × Volán × Malév
Példaképet szeretnék nekik mutatni.	példát szeretnék nekik mutatni × példakép szeretnék lenni
Sosesem szoktam cukrot tenni bele.	soha sem × sosem
Szeretném megköszönni, hogy részt lehet vehettem.	részt vehettem × részt lehetett vennem
Ujjszopás mialatt is előfordul.	miatt × alatt

4. Téves kezdés

Megakadás

A le □ a na □ a dihotikus vizsgálathoz az egyik legnagyobb lapot alkalmazzuk.
A másik sz öö hűtőben van.

Megakadás

A személy együttes le ☐ elhatárolása.
 A szerpentin. Mármint, amit a nyá őö szilveszterkor dobálunk.
 Amit itt most fölül ☐ fölruának...
 Az első és a második helyért hatalmas viha harc van.
 Az igényes nyelvi megfor ☐ őö megfogalmazás...
 Az írott eny ☐ nyelv elsajátítása
 Az összetett és koordinált stuk ☐ struktúrák együtt vannak jelen.
 Döljetek a pada tal ☐ a falnak!
 Egy adott szöveg szó ☐ szóaránya...
 Egyre magas ☐ magabiztosabban őö mozogsz a színpadon.
 Elképesztő vagy, annyira termés ☐ természetesen árad belőled az erő!
 Elkezdtem egy főiskolát, ezt jelenleg is fo ☐ csinálom.
 Én fürödtem már az Északi-tang ☐ tengerben is!
 Ennek még lesznek kövét ☐ ő ☐ következményei
 És megjelent az első szó ☐ szólólemeze.
 Este jól bevar ☐ bevacsoráztam.
 Ez is beadható, ami sin szintén nem egy rossz dolog.
 Ez is bizonyos csel ☐ tevékenységsorozat része.
 Ez is bizonyos csel ☐ tevékenységsorozat része.
 Ez nagyon finom lett ezzel a bezra ☐ besameles mártással!
 Ezen kívül fontos a kiráncó ☐ kizárólagosság.
 Gyorsan be is mulat őö ☐ mutatkozom.
 Ha csak szimplán dilettán ☐ dilettantizmus lenne!
 Hát az megviz ☐ megfizethetetlen.
 Hát most aztán sokat mos ☐ mondtál.
 Hazudott a macska, mert maja ☐ maradt a bajuszán.
 Jó filmet változ ☐ választottál.
 Jól van, ez su ☐ szuper volt.
 Köszönöm, nem kérek, a kaja csak bealsz bealtatna.
 Küldjete egy fé ☐ a fészébúkra egy képet!
 Különböző jelenségekre rel ☐ reflektál...
 Lacika cí ☐ részére
 Le kell fe őö fordulni a körforgalomnál.
 Mama olyan picci ☐ na, kicsi, hogy fél a kutyától.
 ...második része a sz ☐ tanfűzetnek...
 Meg fogja győ... nyerni...
 Megvizsgálom az egé ☐ az összeset.
 Mert a magyar ugye aggli ☐ agglutináló nyelv.
 Mi ang olaszul a konfetti?
 Mindkét fulya ☐ mindkét folyamat ugyanoda fut ki.
 Minél n ☐ több nője volt...
 Most épp nem em ☐ nem elérhető.

Megakadás

Most megyünk a von... buszon.
 Nagyobb menny ☐ létszámú beszélőnél...
 Nagyon kev ☐ kedves volt tőled, jól esett.
 Negatív figurákat fo sorolunk fel.
 Nem is tudom, mit kell né ☐ figyelni
 Nem kell az egész, elég ha a pó ☐ példákat nézzük.
 Nem véletlenül nincsen para ☐ párkapcsolata.
 Nincs Medgyessy-válogatás útbaigaz ☐ útelágazódás nélkül.
 Rá voltam cuppanva a paracs ☐ paradicsomra.
 Segítenél nekem arrébb tolni ezt a komó ☐ szóval kanapét?
 Sikerült dea ☐ diadalmaskodnunk.
 Szerintem is hülyeség ez a mérték ☐ mértek ☐ méregtelenítés.
 Szétvál ☐ szétvágott képeslapot kell összerakni.
 Te honnét szárm ☐ honnét jössz?
 Úgy nézek ki, mint valami ana ☐ anime karakter.
 Uthalhat a beszélő konsz ☐ konzervativizmusára is.
 Üdvözlök mindenkit itt a Gomb ☐ Gömbölyödőben.
 Va ☐ ismerek én is olyan történeteket, vannak olyan családok...
 Venni kéne már szappant ide a fog ☐ csaphoz.

5. „Nyelvem hegyén van” jelenség**Megakadás****Szándékolt közlés**

A ☐ a-val kezdődik ☐ Ankara.
 A digitális irodalmi valami a digitális irodalmi akadémia.
 A göndörítő mi a fene az ☐ hajsütővas, na majd csak kinyögöm.
 Az Andris közben a Karib-szig Karib-tengő Kanári szigeteken csücsül.
 El kéne mosni az izéjét, tudod, mire gondolok, a drótyját.
 Előbb menjünk be az étterembe, nehogy bezárjanak, mert nincs ügyfél, vevő, vagyis vendég.
 én a ő a hát ilyen zeneeszközöket használtam fel a vizsgálatban hangszereket
 Ez a sütemény úgy néz ki mint a ☐ izé ☐ nam bejgli. Dehogyan ☐ molnár laska ☐ nem.
 V van benne ☐ stefánia vagdalt. Vargabéles!

Megakadás**Szándékolt közlés**

Hát csak a metonímiát, meg aa □ mi a más-
sik? Olyan, mint a deizmus, de nem az □
nem deizmus, hanem □ az, amikor előre
meg hátra utalsz a szövegben, meg ki be-
lőle.

Hogy is hívják i □ azt a települést i □ i na,
az a településnév. □ Izsák? Izsák.

Holnap hozd már el aaa hogyhívjád ad-
venti cuccot

Ideadnád az öö SMS-ező készüléket
öömmm a mobiltelefonomat?

Itt mindig van egy csomó kisanya, öröm-
anya, na, kismama.

Megszokta, hogy egyenlő, egyenlő, na hogy
mondják? Nem jut eszembe □ egyenran-
gú útkeresztelkedésben.

Mon □ mi is az □ motiváció.

Nagyon szépen le van írva az e-mail cí □ e-
mail címen, na, az interneten.

Olyan hangja volt mint mint a □ a guppik-
nak, öö Dezsőknek a ketrecben,
deguknak.

Ráesik egy öö hogy hívják öö turbula.

Romparé □ Paroné □ Raponné

Te kérsz kanalat? Nem, villát? Jaj, kést
akartam mondani.

Tekersz csavarsz tekercs öö keversz te?
[kártyát]

Továbbra is egy □ egye □ na, nem jut
eszembe □ egyetemisták.

Tudod, az a nagy zöld gép, nyár végén □na,
mondd már □ jaj □ kombájn!

Vegyük mondjuk aaa □ mi az? Ki neve?

Hogy hívják? □ Gy □ Gy □ Gergő!

Vigyázz, mert jön egy me □ nem, két kere-
ke van, és jármű □ egy motor. Az előbb
metrót akartam mondani.

Voltak ilyen kutatásai öö annak a híres
amerikai szociolingvistának □ l-lel kez-
dődik □ de bosszant □ Labov!

deixis

turbina

6. Perszeveráció

Megakadás	Szándékolt közlés
A Csillagok hábobújá □ bá □ rújában az volt a legjobb, hogy voltak benne robotok.	
A fő különbség a belső monitori öö monitorozásban van.	
A Paulaner's Paulatz is a Bazilikánál van?	Paulaner's Platz
A sarló kalalapács kalapács ősi babiloni szimbólum.	kalapács
Azóta már normanizálódtam.	normalizálódtam
Bocsásson meg, de önök egy húron penbülnék!	pendülnek
Bohókos jelenség.	bohókás
Ebben most mi volt a halmeresztő hal □ hajmeresztő?	
Eltettem a bérletem a tertójába.	tartójába
És akkor ez az ábra fog létrejönni.	létrejönni
Ez nem jelent semlit.	semmit
Felkeltem a felkelt, méltán dicsér a nap.	felkeltem a reggelt
Ha a lé elfő, akkor forró fízzel pótolom.	vízzel
Ha nem tudsz dönteni, akkor dobj fej egy pénzt.	dobj fel
Hol van már a tavalyi hol?	hó
Holnap a gyerekekkel a Hófehérke és a hét hörpét fogjuk olvasni.	törpét
Hupsz, ezt egy kicsit összerentettem!	összerontottam
Itt fogunk kuksulni?	kuksolni
Itt jöttek fel a felmászők.	hegymászők
Jójó, koszi □ köszike.	
Különben is, ez egy hím rémszarvas.	rénszarvas
Még Skype-on szkeresztül sem hallgattam zenét.	keresztül
Megcsikizem a hónananalját, hahaha!	hónalját
Mi sokáig ezt a játszmat játszma □ játszotuk.	
Most kiderül, hogy lesz-e munkám vagy lesz.	vagy nem
Neked sem tetszett, nekem sem tencett	tetszett
Ő állította ötte □ össze.	
Panasz érkezett?	érkezett
Színházba volt vegyem.	jegyem

Megakadás	Szándékolt közlés
Szíveszen segíték.	szívesen
Ti is össze <input type="checkbox"/> össze tudjátok vetni.	
Töltsd még ki azt a kis vörösbort.	vörösbort
Úgy megennék már egy jó kis libamájat libakáposztával.	libakáposztával
Veszek helyette jegesmeciset <input type="checkbox"/> öö macisat.	
Volt fogalomkiegészítéses test <input type="checkbox"/> teszt.	

7. Anticipáció

Megakadás	Szándékolt közlés
...abból a trémakörből...	témakörből
A böldön is bővelkedő életet élnek	a Földön
A saláta jól megy a kumripüréhez.	krumplipüréhez
Az egyik a hallász <input type="checkbox"/> hallás vagy észlelés.	
Az elvileg növeszteli a koleszterint.	növeli
Az internetes sportújságírás <input type="checkbox"/> sportújságírás érdekel a legjobban.	
Az irigység rossz tudajdonság.	tulajdonság
...beszél azokról az elvárosokról, amelyeket a diákok támasztanak.	
Ennyi mindenre nem lehet összefigyelni, már teljesen összezavarodtam.	figyelni
Érezni a tesztszagát.	testszagát
...és a törta <input type="checkbox"/> tartalomösszegzéskor.	
Ez egy szima szivacs.	sima
Ez tök mari, hogy megy a puli a Marsra!	ez tök buli
Ezerkilencszázhatvanhatban.	ezerkilencszázhatvanhatban
Ezt a frusztráltságot próbálják ilyen jellemvonásokkal leplezni.	frusztráltságot
...fél órát maxumum...	maximum
...ha tetán öö netán...	
...hét és fél s <input type="checkbox"/> perces spontán...	
Hint is hideg van.	kint
Hózza a lóbört.	húzza
Két decsi narancslevet kérek öö decit!	
Kezes hazugság.	kegyes
Kézműves díszek díszítik <input type="checkbox"/> táblák	kézműves táblák díszítik
Kicít picit ez a hímozótű, nem?	kicsit
...kiváltott pocen <input type="checkbox"/> potenciálok következtében...	
...küül <input type="checkbox"/> körülbelül azonos mértékben...	

Megakadás	Szándékolt közlés
Megsagolta a saját dezodorját.	megszagolta
Milyen cinos cinos cipőd van.	
Na, mecsek csicsikálni.	megyek
Ne tapizd a majomat!	csajomat
Nekem szerintem ez a reflexem soha nem volt meg.	reflexem
Nem biztos, hogy meg tudod győzni.	Nem biztos, hogy meg tudod győzni
Nem vagy te egy kicsit fulúgos?	flúgos
...nyilván Niki mondta neked, mert csik csak ti magázzátok egymást.	
Ott a vagolban voltam	vagon
Összekaszatoltam □ na, maszatoltam a kezem.	
Perse, sejtettem. Akarom mondani, persze.	
Persze, aztán meg mehetek a sájsebészetre, nem?	szájsebészetre
Rendes volt, átvitt a Pariska öö □ Mariska a Pamacsához	
Szipikus sztereó című szám.	tipikus
Sűrűbben szeretném.	sűrűbben
Tízkor öö tízkor kezdődik a film.	
Tőled tollerálom.	tolerálom
Zsöld zsenília pulcsi.	zöld

8. Metatézis

Megakadás	Szándékolt közlés
...a mező madarai és az ég liliomai...	az ég madarai és a mező liliomai
A lányok beletörték a szekrényt a kulcsba.	
Mi? A kulcsot a szekrénybe.	
A tószőt □ szótót előhívom.	
Aha, igen, a tuján fékszel a rigó.	fészkel
Áin és Kábel, nem □ nem, Káin és Ábel.	
...az ojal □ olaj...	
Bogyóolajokat kell enni, az jót tesz, igen, azt mondják, az olajbogyó nagyon egészséges.	
Csényi kínálni tortát.	kéne csinálni
Csukd be a hideget, mert bejön az ablak!	Csukd be az ablakot, mert bejön a hideg!

Megakadás	Szándékolt közlés
És idővel jó és jó idővel azonosították a szavakat.	
És ott volt a szörp a zsebed farmerjében.	farmered zsebében
Ez a feteke nekem egyáltalán nem tetszik.	fekete
Ez nem a sokászos reggeli híradó.	szokásos
Fát lehet hágni a vágukon.	vágni a hátukon
...felek kel □ kelek fel...	
Hallom a fejeket a hangomban.	hangokat a fejemben
Hangsúlyosak még a névszámok □ ah, számnevek is.	
Hihetetlen, hogy mindenki azt az ugyane-gyet akarja!	ugyanazt az egyet
Igazán praffáns kis előadást hoztál össze!	frappáns
Jól visszazártam a paplant. Hmmm. Vissza-ráztam.	
Kérem, záródjanak, az ajtók vigyáznak!	vigyázzanak, az ajtók záródnak
Leírták, hogy a tószövek öö szótövek ho-gyan tárolódnak.	
...leporolni a tört...	letörölni a port
Lesz két kiló raback öö barack.	
Ma megyek majd be az Apácia □ Apácaiba	
Maradj inkább le tenn!	te lenn
Mátyás király állandóan fejlesztette hadóha-ját, 1476-ban már 364 hajója volt.	hajóhadát
Meg a szupre □ szuperbruttó is.	
Még kennem vell kaját holnapra.	vennem kell
Még mindenképpen nézzük meg a széltacsornát!	szélcsatornát
Még nem szívta be a krém a kezemet.	a kezem a krémet
Megmakacsolta magát, hogy ő nem csinál hülyéből senkit vagy senkiből hülyét.	
Mindig azt hiszem, hogy ez a lekrám ne-kem szól.	reklám
Nem beszéltem anyádról ilyenrel.	ilyenről anyáddal
Nem csak a gyerekeim, hanem az unokáim is kezdenek már fölcseperedni.	fölcseperedni
Nem mindent összenek önte.	öntenek össze
Nyugi! Mélyegezz léleket!	Lélegezz mélyeket!
...porog-förög...	pörög-forog
Sükszükszerű, hogy megtegyük a lépéseket.	szükségszerű
Ú, odavagyok a teáért, szerintetek lehet valahol venni cseraterjét?	teacserjét

Megakadás	Szándékolt közlés
Van kép a bőrödön <input type="checkbox"/> vagy van bőr a képeden	
Várj, csak felhúzom a lábamat a csizmámra!	a csizmámat a lábamra

9. Egyszerű nyelvbotlás

Megakadás	Szándékolt közlés
A hírek szerint a korány <input type="checkbox"/> kormány...	
A implantáció eszközrendszere...	az
A Kunigunya útján lakom öö Kunigunda.	
A múltbeli csekedet <input type="checkbox"/> cselekedetet nem nézik már ilyenkor	
Az az alapvető fusztráltság, ami benne van azokban az emberekben, az hihetetlen!	frusztráltság
Az egyik hallász <input type="checkbox"/> na, de jó vagyok <input type="checkbox"/> hallás...	
Az eladó az előzetes nyelvészeti és pszicholingvisztikai tanulmányaiból eredően...	előadó
Az itt megjelent szavak már egy <input type="checkbox"/> az itt megjelent szavaknak is már egy...	
Az Ogli Dí <input type="checkbox"/> Ogli Dzsí...	(Ogli G)
Aztán beletöltjük a pisokótába.	piskótába
Belenyomon a fejedet <input type="checkbox"/> belenyomom	
De miért vekszel?	veszekszel
Egész nap csak ordít, mint egy féreg, mint egy fába szorult féreg.	
Első körben akkor próbáljuk meg feleveníteni a múlt órán tanultakat!	feleveníteni
...és ezt svábbal írjuk.	svával
Holnap jössz korlyázni?	Holnap jössz korcsolyázni?
Ingyenes nyíl hét <input type="checkbox"/> nyílt hét áll az érdeklődők rendelkezésére.	
Itt a szövegterkezet öö tervezet.	
Itt egy gis kocsmá.	kis kocsmá
Józnakak tűndek <input type="checkbox"/> tűntek.	
József körút Baross uccsa sarokra menénk.	Baross utca
két háccal arrébb <input type="checkbox"/> két házzal arrébb	
Korlátok közé sorul ö szorul.	
...koznyelvi <input type="checkbox"/> köznyelvi alakok...	

Megakadás	Szándékolt közlés
Közben próbálkozok a beszével <input type="checkbox"/> beszéd- del.	
Közgazdaszságtant végeztem.	közgazdaságtant
Láthatjuk, hogy ez egy életspecifikus <input type="checkbox"/> életkorszakspecifikus jelenség.	
Leginkább a metatézis bizonyítás típusai- hoz.	bizonyos
Manyciul meg úgy van...	manysiul
Milyen korcos vagyok!	kócos
Minden téma uránt mély érdeklődésük van.	iránt
Most nyuszalt bele nagyon a lábamba!	nyilallt
Na meszéld már el!	meséld
Ne már, teljesen tönkre fogunk menni, any- nyi itt a pesszicizmus.	pesszimizmus
Nem válatott <input type="checkbox"/> váratott magára sokáig.	
Ott van az a rengeteg fogkefa.	fogkefe
szép ajándok	ajándék
Szókratész célja nem az volt, hogy a halálbüntést <input type="checkbox"/> büntetést elkerülje.	
Ugye tudod, hogy nagyon nem bírom a rózasínt!	rózsaszínt
Utána leadom a sportán <input type="checkbox"/> a portán.	
Zsófi apukuktúrán van.	akupunktúrán

10. Több típusba sorolható jelenségek

Megakadás	Szándékolt közlés
A Baba Negrában volt a koncert.	Barba Negrában
A beszédtechnika órán ilyen órákat <input type="checkbox"/> órá- kat, feladatokat, elnézést.	
A birtokoshoz tartozó kritikumon értjük...	klitikumon
A disz <input type="checkbox"/> diszkurz <input type="checkbox"/> diskurzusban van szerepe.	
A film után lesz egy sajtótájékoztató.	sajtótájékoztató
A harmadik évad tájékán már nem neki szurkoltam.	tájékán
A hatos <input type="checkbox"/> hatás <input type="checkbox"/> a gyakorlott hatás felejt- hetetlen volt.	
A leggyakorabb említett ezek közül...	a leggyakrabban
A lírai kö <input type="checkbox"/> versekben, költeményekben...	
A mosogatógépbe <input type="checkbox"/> a mosogatógéphez <input type="checkbox"/> a mosógéphez is lehet ilyet venni.	

Megakadás	Szándékolt közlés
A papagáj ilyeneket mond, hogy „beleme- gyek a veredbe!”	véredbe
A szabályok éttermében őö annyit beszél- tünk az indiai ételekről...	értelmében
A téma detonációja ő denotációja...	
Ahhoz, hogy jól jejenjen meg, a rossz szö- veget ki kellett javítani.	jelenjen
Ahogy az öblös orgánumból □ orgánumból □ orgánumomból is látszik.	
Aki csinálta a kasztot mi az izé □ színész- válogatást, nagyon okos.	castingot
Akkor most egy kör békeügetés.	békaügetés
Akkor most mondok egy nagy bölcsességet megint: nem a győzelem a részvét, ha- nem a fontos!	nem a győzelem, hanem a rész- vétel a fontos
Arról se tud senki lelevelni, hogy...	lenevelni
Asztalitiz □ asztaliteniszezem már négy éve.	
Az ember lépdeléséhez hasonlatos érzést hoz lépre.	hoz létre
Az igekötő helyes és az ige nem, vagy az igekötő helyes és az igekötő helyes.	igekötő nem
Az összes ilyen trutymó felkerül az aszmoszférába, aztán majd megnézhetjük magunkat.	atmoszférába
Azért csináljuk, mert montos. Izé, fontos, hogy pontos.	
Azon rőfögtünk, hogy őö röhögtünk...	
Azt kell mondjam, elég jó kérőd van!	kérőd
Aztán a vezetőség, vagyis hát a király úgy döntött, hogy vezessenek aranypénzt.	veressenek
Benne venük □ bennük van a zene.	
Dávid itt hagyta a fél kár kesztyűjét.	pár
De most akkor tényleg tudsz nekem ilyen pásos sálat kötni?	sávós
De nem használhatom bármilyen helyzet- ben a <i>viszont</i> mindenhol <i>viszontot</i>	<i>viszont</i> szót
Dugó volt a Margot-hídon.	Margit-hídon
Ebben a nézetben van két tucat mással- hangzó	négyzetben
Egy lencsefőzelékbe bele fogok pakolni □ bele fogom pakolni az arcomba.	lencsefőzeléket

Megakadás	Szándékolt közlés
Egy nagyon sajátos helyzetet be <input type="checkbox"/> egy nagyon sajátos helyzetet mutat be.	
Elég sok gyökérrel rendelkezett.	gyökérrel
Én nem szeretem, ha szab <input type="checkbox"/> ha szívességet adnak.	tesznek
Én nem vagyok másnap soha.	másnapos
Engem ábrá <input type="checkbox"/> belőlem ábrándultál ki	
Erős Pistát ittél, sört ittál.	ettél
És itt jössz te a gépbe.	képbe
...és majd sliccelünk.	sniccerезünk
...esősen szorongó kisfiú esetén...	erősen
Ez az adag tejszín feleennyibe is elég.	feleennyibe
Ez egy nagyon fontos megbeszélgetés lesz.	megbeszélés
Ez ilyen se veled, se vele kapcsolat.	se nélküled
Ez nagyon fakhagymás.	fokhagymás
Ez olyan, hogy sok ember nem tud megenni egy elefánt!	elefántot
Fekete bika kopog a fekete patika övein...elszúrtam.	a fekete patika kövein
Figyu, kicsit most már elegendem van, hogy mindig engem kötekedsz!	velem kötekedsz
Gyakoriak az antipicá <input type="checkbox"/> anticipációk.	
Gyerekek, hát ez igazán kedves, nagyon avajosak vagytok, komolyan.	aranyosak
Gyomrott a korgóm, vagyis gyomrott <input type="checkbox"/> most mi van?	korgott a gyomrom
Gyorsan kész volt a tészta: fél hatra már össze volt gyűrve.	össze volt gyúrva
Ha továbbra is így viselkedsz, az lesz a vége, hogy ideültetlek az első padba, az orromba!	az orrom elé
Ha valaki elvesztette a sulykot...	elvetette
Hányszor kelsz?	hányszor
Hátrarúgott és mögötte az egész sör borult <input type="checkbox"/> őö <input type="checkbox"/> a sör borult. [sörözés közben]	
Hatvanötfolé fonéma.	Hatvanötféle fonéma.
Hogy egy aktualitassa <input type="checkbox"/> aktualitással kezdjem...	
Hogyan vannak a topok <input type="checkbox"/> toposzok szerkesztve?	
Hosszúval kell inni?	írni
Hova van a papucsom?	hol

Megakadás	Szándékolt közlés
Humor és Magor	Hunor
Ín □ én így nem tudok beszélni.	fűre
Itt most miért van földre lépni tilos?	hasonlít a másakra × hasonló a
Itt tudjuk kiszámolni, hogy az egyik érték mennyire hasonló a másakra.	másikhoz
Ja, azért tetszik neked, mert tizenkét megepixeles.	megapixeles
Jaj, nem lehetne, hogy ne a fec foci legyen a téma?	
Jobban próbáljuk magunkat mono mono monitorizálni.	monitorozni
Kakukkosfű legyen a leves vagy inkább ne? ...kapcsam köl □ kaptam kölcsön a könyvet.	kakukkfűves
Kapirgálhatnék a szeméthonbon.	szemétdombon
Kár, hogy nem kóstoltad meg azt a csömörszét ... szömörleszt ... szömörleszt.	
Két pa két szék közt a gödörben.	két szék közt a pad alá
Két típusú ember van, aki sssz összességében látja a dolgokat és aki pic □ kicsi részleteiben.	
Kevés a vér □ a vas a véremben.	
Könnyen elfordulhat máskor is.	előfordulhat
Látod azt a falat?	fa halat
Legkönnyen ott fog elszakadni.	legkönnyebben
Légyszi, gyűjts tüzet!	Légyszi, gyűjts tüzet!
Ma reggel kivételesen teljesen friss és fett vagyok, úgyhogy semmi sem állíthat meg!	fitt
Már keverem a leveket... leveket, neveket!	
Már megint leköpöcselted a ruhád!	lecsöpögtetted
Mások a Kimilindzsárón vannak, én meg itt.	Kilimandzsárón
Meg kell nézni fogalom között □ fogalomkörök között a kapcsolatot.	
Megkapta-e már az ifluenza elleni védőoltást?	influenza
Mert ki kell, hogy irtsam ezeket a szavak □ számakat.	számokat
Mert lehetőségük lesz a tö az orvos szakma többi területét is felfrissíteni.	

Megakadás	Szándékolt közlés
Mesélte, hogy hantival fótballozott.	Hontival futballozott
Milyen csiná <input type="checkbox"/> milyen teát csináljak?	
Nahát, hogy iszi a whiskeyt!	issza
Negyvenhatezer forint és tettől többet nem.	ettől
Nem akar fotofinte foteszi fotoszintetizálni.	
Nem akarsz gasztrobogot, bogot?! blogot indítani?	
Nem ez az első greenbrox kamerapróbám.	greenbox
Nem fázom, mert bérelt kabátban vagyok.	bélelt
Nem felelt meg az eljárásoknak <input type="checkbox"/> elvárásoknak.	
Nem mindegyik be van sorolva	van besorolva
Néma vagy kitöltött szután <input type="checkbox"/> szünet után...	
Nézd meg a szem <input type="checkbox"/> kezemet!	
...nyelv és fejleszt <input type="checkbox"/> nyelv és beszédfejlesztő...	
Nyilvános a diktori doktori szigorlat.	
Nyolc volt a terpeszkedő szervezet öö <input type="checkbox"/> szerkezet.	
Ő a részéve vá <input type="checkbox"/> akar l <input type="checkbox"/> öö ennek lenni.	
Ötletem sincs, hol tettem.	hová
Remélem, egy év múlva egy komoly Már <input type="checkbox"/> komoly Márk fog visszaköszönni innen.	
...részlegesen öltözködöl.	rétegesen öltözködsz
Se te se sértegecsél!	Te ne sértegecsél!
Segítetsz, ha tudsz!	segíthetsz
Sosasem értettem, minek kell ez.	sohasem
Süsögesd addig meg.	sütögesd
Szép ilyenkor a táj, mert olyan értetlen. Értetlen?! Érintetlen.	
Szeretnének egy náluk nagyobb karand kaland része lenni.	
Szerintem ez az összeolvasás jelensége.	összeolvadás
Szerintem ő mindig békázik mármint békávézik.	
Szomorúan fogjuk észet az esztét zární.	ezt az estét
Szögesen nem értettek egymással szemben	szögesen ellentmondtak egymásnak × nem értettek egymással egyet
öő szögesen nem értettek egymással el-lent.	szünet a szóban
Szünet a szó jelenség..	
Találtam Weöres Sándor-se <input type="checkbox"/> verseket.	
Tömve volt a dudvar is.	udvar is

Megakadás**Szándékolt közlés**

Ugye felírtátok, hogy tószövek öö szótövek?

Van egy tudatos kontroll a kíséretben ☐ kíséretben résztvevőknél.

Vannak olyan emberek, akikhez annyira hozzánőtt egy-egy testrésük ☐ az is lehetett volna, de inkább testrésük.

Vonj egy visszanyomást ☐ nyomj egy visszavonást!

A MAGYAR ÉS A NEMZETKÖZI TUDOMÁNYOS ÉLET ESEMÉNYEI

Björn E. Lindblom 80 éves

Björn Lindblom sokak számára talán a beszédtudomány nagy alakjaként ismert, mások feltehetőleg a fonetikai és fonológia közti merev határok fellazítására tett erőfeszítése miatt, emergens fonológiai elméletéért ismerik.

Lindblom 1934-ben született Stockholmban. Diplomáit angolból és romanisztikából szerezte a Stockholmi Egyetemen, majd 1968-ban doktorált fonetikából Lundban. Ekkoriban a svéd Királyi Műszaki Intézetben (KTH) gyakor-nokoskodott a fonetika doyenje, Gunnar Fant mellett, később pedig az akusztikus Kenneth N. Stevens vezetése alatt az MIT Beszédkommunikációs Osztályán is tevékenykedett. A Stockholmi Egyetem professzoraként a mai napig aktív kutatómunkát végez, vendégprofesszori előadásai Amerikában (a Texasi Egyetemen) és Svédországban (a KTH-n és a Stockholmi Egyetemen) egyaránt hallhatók. Számos díjat és kitüntetést tudhat magáénak, de szakmai nagyrabecsülését legjobban talán a kollégái és tanítványai által szerkesztett, a hatvanadik születésnapja alkalmából kiadott *Phonetica* folyóirat különszáma Engstrand–Kohler eds. 1994. *Speaker, listener and the child*), illetve a hatvanötödik születésnapjára összehívott nemzetközi szimpózium (*Emergence and Adaptation: Speech Communication and Language Development*) és az abból készült tanulmánykötet példázza (Diehl et al. eds. 2000).

Lindblom tanulmányai között nagy számban találunk kísérleti fonetikai munkákat és a kísérleti eredményekre támaszkodó elméleti írásokat is. Pályája kezdetén (az 1960-as évektől) Fant nyomdokain a beszédhangok, különösen a magánhangzók ejtésének artikulációs és akusztikai megvalósulását kutatta. A korabeli legfejlettebb technológiai vívmányokkal (röntgenográfiával, videofelvételekkel, később mágneses rezonancián alapuló képalkotó eljárásokkal) végzett elemzései a **beszéd artikulációs modellezésére** tett nagy ívű projektum részei voltak. A kutatás célja az **APEX** artikulációs modell és az arra épülő szoftver megalkotása volt. A toldalékcső (beszélőspecifikus) csőmodelljének elkészítéséhez Lindblom és munkatársai hatalmas mennyiségű képanyagot dolgoztak fel. A felvételeken az ajkak, a nyelvcsúcs, a nyelvtest, az állkapocs és a gége helyzetét mérték, ezeket felhasználva kiszámították a toldalékcső egyes szakaszainak átmérőjét, ezekből pedig meghatározták az átviteli karakterisztikát. Az APEX szoftver segítségével az artikulációs konfiguráció paramétereit (az állkapocs nyitásszögét és a nyelv helyzetét) változtat-

hatjuk, ami alapján a formánsértékek kinyerésén keresztül a program szintetizálja a beszédhangokat (l. pl. Stark et al. 1996).

A beszédhangok artikulációs és akusztikai megvalósulásának kapcsolata a beszédpercepcióval a fonetikai kutatások máig kimerítetlen forrását képezik. A beszélők közötti és az egy beszélő ejtésén belüli variabilitás miatt ugyanis kérdés, hogy vajon hogyan lehetséges az, hogy még az egymástól sokszor igen jelentős mértékben eltérő beszédhangokat is képesek vagyunk azonosaként felismerni. A kutatók abban egyetértenek, hogy a felismerési feladat sikerességéhez a változatosság ellenére is állandóan jelen lévő invariáns jegyek megléte (és azok felismerésének képessége) lenne szükséges. Eltérő elgondolások léteznek azonban arra nézvést, hogy milyen természetűek ezek az invariáns jegyek. A kutatók egy csoportja ezeket a jegyeket az akusztikai megvalósulásban (l. Stevens és Blumstein munkáit), mások pedig az artikulációban (l. a motoros elmélet képviselőinek munkáit, pl. Liberman és Mattingly írásai) keresik. Lindblom saját kutatási eredményeire támaszkodva azonban mindkét nézőponttal szembehelyezkedik. Úgy véli, az állandóság valójában nem az akusztikai vagy artikulációs jelben, hanem a hallgatóban keresendő, ugyanis az állandóság nem akusztikai vagy artikulációs jellemző, hanem a beszéd reprezentációjának tulajdonsága. Ezek a gondolatok képezik a kutató talán leghíresebb munkájának, a **H&H elméletnek** az alapjait, melyet Lindblom 1990-ben publikált. A **H&H** név a *hyper-* és *hypospeech* kifejezések rövidítéséből áll elő, melyek a fonetikai információban gazdag és az információhiányos jel koncepcióit fedik (egy kontinuum két végpontjaként). Lindblom elmélete szerint a beszélő a beszéd interakciós helyzetében az információgazdagság és információhiány skálája mentén rendelkezésre álló lehetséges megvalósítási formák közül választ annak fényében, hogy mit feltételez a hallgató számára (a jeltől függetlenül) elérhető információról. Lindblom szerint a beszéd leképezésére szolgáló komplex reprezentációinkat azon jelekből konstruáljuk a nyelvsajátítás során, melyeket a beszéd fonetikai és nem fonetikai jegyeiből ismerünk fel annak rendszerszerű változatossága révén. Ebből pedig az következik, hogy az állandóság a reprezentáció tulajdonsága, és (paradoxnak tűnő módon) valójában a variabilitásban rejlik. A variabilitás tehát nem a beszéd folyamat másodlagos vagy mellékes, hanem lényegi, esszenciális jellemzője (l. pl. Lindblom 1990, Lindblom et al. 1992).

A változatosságnak mint a beszéd esszenciális tulajdonságának gondolata motiválhatja közvetett módon Lindblom érdeklődését a koartikulációs folyamatok iránt is. A koartikuláció jelentése 'együttes artikuláció', és a beszédhangok beszédfolyambeli egymás utáni ejtéséből fakadó egymásra hatására vonatkozik. Az egymásra hatás mikéntje számos ponton nyitott kérdés, ahogy bio-fiziológiai oka, eredete sem triviális. Lindblom koartikulációs elmélete ez utóbbival, tehát a **koartikuláció motoros működésének eredetével** foglalkozik. Az elmélet szerint a koartikuláció, azaz az egyes beszédhangok ejtéséhez szükséges precíziós mozgások időbeli és térbeli átfedése a beszéd ejtése

közben zajló kétféle minőségű motoros működés interferenciájából fakad. Azon mozgások nagy részét, melyek az ember mindennapi vegetatív működését (rágás, lélegzés, járás) irányítják, nagyfokú automatizáció, ezzel összefüggésben pedig nagymértékű ritmikusság és viszonylagos gyorsaság jellemzi: ezek „elemi rezgőmozgások”. Azon mozgásokhoz azonban, melyek precíziós manőverezést igényelnek (discrete target control, pl. létrára mászás, tárgyak megfogása), aktívabb, tudatos kontroll szükséges, éppen ezért működésük is lassabb és kevésbé ritmikus. A beszédképzés folyamatában, írja Lindblom, ugyan az állkapocs mozgását az automatizált kontroll vezérli (nem véletlen, hogy a beszéd és a rágás igen hasonló mozgásmintázatot mutat), tehát gyors és ritmikus, a hangzójtéshez szükséges további precíziós beállításokat jóval nagyobb mértékű felügyeletet igénylő mozgások segítségével, így valamelyest lassabban is végezzük. Az egyes artikulációs gesztusok időbeli és térbeli átfedése tehát, Lindblom elmélete szerint, ennek a kétféle mozgási működésnek az időzítési sajátosságaiból fakad. Az elmélet szerint az állkapocsmozgás ciklikus természete a szótag jelenségére és létrejöttére is egyfajta (emergens) magyarázatot ad (l. Lindblom–MacNeilage 2011).

Lindblom több írásában is foglalkozik az **emergens fonológia** elképzelésével. Az emergens fonológia sokban épít a H&H-ban már megfogalmazott alapelvekre, melyek szerint a beszédértés rendszerében nem a jel feldolgozásának folyamata, hanem a tárolás az összetett (szemben állva azokkal a feltételezésekkel, melyek gazdaságossági elvekre hivatkozva a feldolgozás komplexitását és a tárolás egyszerűségét feltételezik). Az emergens elmélet javaslata szerint ugyanis nem kis mennyiségű és lecsupaszított (invariáns) adatot raktározunk, melyeket bonyolult elemzés során nyerünk ki a beszédjelből, hanem éppen ellenkezőleg, komplex és sokszor redundáns információstruktúrákat tárolunk, melyeket egyszerű és robosztus jelfeldolgozással nyerünk. Lindblom úgy véli, a nyelvelsajátítás során a gyermekek a tárolt komplex spektrális információhalmazokból alakítják ki későbbi nyelvi rendszerük perceptuális kategóriáit klaszterezés segítségével, azok tehát kumulatív fonetikai tapasztalataink nyomán „emelkednek ki”. A beszédprodukció fejlődésével kapcsolatban Lindblom úgy véli, a gyermek első, imitációs célú artikulációs mozgásai az összes többi mozgáshoz hasonlóan a minimális energia kritériumát követik, tehát az első hangsorok produkciójában is a „legkisebb költségű” elemi rezgőmozgás lesz jellemző: az állkapocs nyitogatása semleges toldalékcso-konfigurációval (*baba, papa*). Azt a perceptuo-motoros kapcsolatot pedig, mely a perceptuális kategóriák és az említett véletlenszerű artikulációs működések között jön létre, a környezeti megerősítés biztosítja. Az elmélet szerint az ejtett hangok tárának fejlődését az artikulációs csatorna struktúráinak részleges elkülönülése segíti, a gyermek számára ugyanis azok a hangok lesznek elsajátíthatóak, melyek egyes „részei” (azaz az artikulációjához szükséges egyes elemek) már benne vannak a gyermek artikulációs repertoárjában. A fonémakészlet tehát az ilyen módon „átfedésben lévő” hangok irányá-

ba bővíülhet. Ám ugyanezen inherens szegmentálás teszi lehetővé a fonetikai gesztusok reprezentációjának kialakulását is. Látható tehát, hogy az emergens fonológiai elmélet azon elméletek közé tartozik, melyek a nyelvet nem elemekből és szabályokból építkező, hanem analógián, szabályáltalánosításon alapuló rendszernek tekintik, és nem feltételeznek veleszületett és nyelvspecifikus mechanizmusokat, hanem a nyelvet az általános kognitív képességek részének, az általános megismerésen és annak kategóriáin alapuló rendszernek tartják (l. Lindblom 1999).

Lindblom egy sokat vizsgált témája a világ nyelveinek hangzókészlete is, ami részben szintén az emergens fonológia melletti elköteleződésével magyarázható. Ha ugyanis a fonológia valóban emergens, akkor a világ nyelveiben kialakult hangzók sem lehetnek véletlenszerűek, sokkal valószínűbb, hogy a motorikus működések elvei által motiváltak – vizsgálatuk tehát segíthet jobban megismerni a fonológia kialakulásának univerzális folyamatát (l. pl. Lindblom 1999).

Bár a fenti összefoglalásban csak szemezgetni volt alkalmunk Björn Lindblom életművéből, mégis talán sikerült megvilágítanunk, milyen változatos és példaértékű munkásság az övé. Ötvözte a kísérletes munkát az elméletalkotással, és sokat tett azért, hogy a fonetika tudománya ne csak leíró, hanem magyarázó értékű is legyen. Szívből kívánjuk, hogy még hosszú éveken át folytassa ezt az eredményekben gazdag munkát.

Björn Lindblom néhány munkája:

- Hura, Susan – Lindblom, Björn – Diehl, Randy 1992. On the role of perception in shaping phonological assimilation rules. *Language and Speech* 35/1–2. 59–72.
- Lindblom, Björn 1990. Explaining phonetic variation: A sketch of the H&H theory. In Hardcastle, William – Marchal, Alain (eds): *Speech production and speech modeling*. Kluwer, Dordrecht. 403–439.
- Lindblom, Björn 1996. Role of articulation in speech perception: Clues from production. *Journal of the Acoustical Society of America* 99/3. 1683–1692.
- Lindblom, Björn 1999. Emergent phonology. In: *Proceedings of the Twenty-fifth Annual Meeting of the Berkeley Linguistics Society*. University of California, Berkeley. 195–209.
- Lindblom, Björn – Brownlee, S. – Davis, Barbara – Moon, Seung-Jae 1992. Speech transforms. *Speech Communication* 11. 357–368.
- Lindblom, Björn – MacNeilage, Peter F. 2011. Coarticulation: A universal phonetic phenomenon with roots in deep time. *Fonetik 2011*. 41–44.
- Lindblom, Björn – Sundberg, Johan 1991. Acoustical consequences of lip, tongue, jaw and larynx movement. *Journal of the Acoustical Society of America* 50. 1166–1179.
- Stark, Johan – Lindblom, Björn – Sundberg, Johan 1996. APEX – an articulatory synthesis model for experimental and computational studies of speech production. *Fonetik 96. TMH-QPSR* 1996/2. 45–48.

Sundberg, Johan – Lindblom, Björn – Liljencrants, Johan 1992. Formant frequency estimates for abruptly changing area functions: A comparison between calculations and measurements. *Journal of the Acoustical Society of America* 91/6. 3478–3482.

A Björn Lindblom tiszteletére kiadott kötetek:

Diehl, Randy – Engstrand, Olle – Kingston, John – Kohler, Klaus (eds.) 2000. Emergence and adaptaion: Speech communication and language development. *Phonetica* 57/2–4.

Engstrand, Olle – Kohler, Klaus (eds.) 1994. Speaker, listener and the child. *Phonetica* 51/1–3.

Deme Andrea

A KÖTET SZERZŐI

Név	Intézmény	E-mail
Abari Kálmán	Debreceni Egyetem Pszichológiai Intézet	abari.kalman@arts.unideb.hu
Bartalis Mátyás	BME Távközlési és Média-informatikai Tanszék	bartalis@tmit.bme.hu
Beke András	MTA Nyelvtudományi Intézet	beke.andras@nytud.mta.hu
Bóna Judit	ELTE Fonetikai Tanszék	bona.judit@btk.elte.hu
Csapó Tamás Gábor	BME Távközlési és Média-informatikai Tanszék	csapot@tmit.bmea.hu
Deme Andrea	MTA Nyelvtudományi Intézet	deme.andrea@nytud.mta.hu
Gocsál Ákos	PTE Pollack Mihály Műszaki és Informatikai Kar és MTA Nyelvtudományi Intézet	gocsal@gmail.com
Gósy Mária	MTA Nyelvtudományi Intézet	gosy.maria@nytud.mta.hu
Grácz Tekla Etelka	MTA Nyelvtudományi Intézet	graczi.tekla.etelka@nytud.mta.hu
Horváth Viktória	MTA Nyelvtudományi Intézet	horvath.viktoria@nytud.mta.hu
Imre Angéla	ELTE Fonetikai Tanszék	imre.angela@btk.elte.hu
Kiss Jenő	ELTE Magyar Nyelvtörténeti, Szociolingvisztikai és Dialektológiai Tanszék	kiss.jeno@btk.elte.hu
Kohári Anna	ELTE Fonetikai Tanszék	koharianna@gmail.com
Markó Alexandra	ELTE Fonetikai Tanszék	marko.alexandra@btk.elte.hu
Mátyus Kinga	MTA Nyelvtudományi Intézet	matyus.kinga@nytud.mta.hu
Németh Géza	BME Távközlési és Média-informatikai Tanszék	nemeth@tmit.bme.hu
Neuberger Tilda	MTA Nyelvtudományi Intézet	neuberger.tilda@nytud.mta.hu
Olaszy Gábor	BME Távközlési és Média-informatikai Tanszék	olaszy@tmit.bme.hu

Név	Intézmény	E-mail
Orosz György	PPKE, Információs Technológiai és Bionikai Kar és MTA-PPKE Ma- gyar Nyelvtechnológiai Kutatócsoport	oroszgy@itk.ppke.hu
Szántó Anna	ELTE Fonetikai Tanszék	szantoanna1984@gmail.com
Tóth Andrea	ELTE Fonetikai Tanszék	toandi@mailbox.hu
Váradi Viola	ELTE Fonetikai Tanszék	varadi.viola@btk.elte.hu